The Epigenetic Regulation of the EGF-receptor Ligands Amphiregulin and Epiregulin and Its Impact on the Outcome of EGFR-targeted Therapies.

D i s s e r t a t i o n
zur Erlangung des akademischen Grades
d o c t o r r e r u m n a t u r a l i u m
(Dr. rer. nat.)
im Fach Biologie
eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät I
der Humboldt-Universität zu Berlin
von

Dipl. Biochem. Felix Udo Bormann

Präsident der Humboldt-Universität zu Berlin
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät I
Prof. Stefan Hecht, Ph.D.

Gutachter/innen: 1. Prof. Dr. C. Sers
2. Prof. Dr. N. Blüthgen
3. Prof. Dr. A. Ehrenhofer-Murray

Tag der mündlichen Prüfung: 24.04.2014

Die vorliegende Arbeit wurde vom 01.05.2009 bis 19.11.2013 im Institut für Pathologie an der Charité Berlin in Berlin-Mitte in der Arbeitsgruppe von Frau Prof. Dr. Christine Sers angefertigt.

ACKNOWLEDGEMENTS

This work would never have been possible without the support of Professor Dr. Christine Sers. She gave me a topic which combined epigenetics with systems biology and might have a strong influence in cancer research. She gave me enough freedom in designing and performing experiments, but also guided me through the complete work, that I could not loose my focus.

Then I'd like to thank all lab members of the lab of molecular tumorpathology for supporting me in any aspects, starting from answering simple questions, continuing with helping me in performing experiments and ending with being good friends. Thanks goes to Christina Kuznia, Dirk Schumacher, Sha Liu and Paula Medina. Special thanks goes to Natalia Kuhn and Stephan Bartels, who actively contributed to some of the findings of this work.

I specially thank our MTA Cornelia Gieseler. She helped me when I was upset with experimental designs and experimental results and showed me, that some of the experiments are not as difficult as I expected them to be. She was also always a contact person for any problem which arose during the work.

I would also thank the members of the ColoNET-consortium, especially Maria Rivera and Sascha Tierling, who performed the Xenograft-experiments and the methylation analysis experiments which I was not able to do, but which were very important for the work presented here.

My special thank goes to my friends and family who supported me mentally and never stopped to encourage me, especially my grandparents, my parents and my brother Tobias.

The last person, I would like to thank is Johanna. At any time of this work, you were there to encourage me and to listen to my problems. You helped me through the most depressive moments by showing me the positive site of the coin, for example by suggesting solutions to any problem and giving me motivation. Without you, this time of my life would have been much harder. Thank you.

Contents

Abbreviations

AREG	Amphiregulin
bp	basepairs (DNA)
bs-cDNA	bisulfite-converted cDNA
bs-DNA	bisulfite-converted DNA
BTC	Betacellulin
cDNA	complementary DNA
CIMP	CpG island methylator phenotype
CIN	chromosomal instability phenotype
CpG	Cytosine-Guanine-dinucleotide
CRC	colorectal cancer
Ct	cycle of treshold
CTCF	CCCTC-binding factor
CTCFL	CCCTC-binding factor like
DAC	Decitabine (5-aza-2'deoxicytidine)
DNA	deoxyribonucleic acid
dNTP	deoxyribonucletide
DNMT	DNA methyltransferase
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EMA	European Medicines Agency
EREG	Epiregulin
EPGN	Epigen
FAM	fuorescent dye (6-Carboxyl-Fluorescein)
FDA	Food and Drug Administration
HAT	histone acetyltransferase
HDAC	histone deacetylase
HDACi	HDAC inhibitor
hg	human genome
LOH	loss-of-heterogeneity phenotype
kb	kilobases (DNA)
MAPK	Mitogen-activated protein kinase
MCS	multiple cloning site
MI	methylation index
mm	mus musculus genome
mRNA	messenger RNA
MSI	microsatellite instability phenotype
msSNuPE	methylation-sensitive single nucleotide primer extension
NSCLC	non-small-cell lung carcinoma

PCR	molymerase chain reaction
PI3K	Phosphoinositide 3-kinase
RNA	ribonucleic acid
TGF-α	Transforming growth factor α
VIC	fuorescent dye from ABI
ZBTB4	Zinc finger and BTB domain containing protein 4
ZBTB33	Zinc finger and BTB domain containing protein 33

List of Figures

List of Tables

1 Introduction

1.1 Colorectal cancer

In 2008, colorectal cancer was the second most common cancer type with 14.3 % new cases among all male and 13.5 % new cases among all female cancer patients in Germany.[1] More frequent was only prostate cancer with 25.7 % in males and breast cancer with 32.1 % in females. With a mortality rate of 11.8 % among all male and 13 % among all female cancer induced deaths, colorectal cancer was also the second major cause of cancer induced mortality. The life-time risk to develop colorectal cancer was 7.5 % for males and 6.1 % for females. Additionally, the life-time risk to die from colorectal cancer was 3.2 % for male patients and 2.7 % for female patients. Dietary factors like high consumption of red meat, unsaturated fatty acids and alcohol, as well as inherited and somatic mutations, are some of the risk factors for the development of colorectal cancer.[2,3,4]

To describe the different stages of spontaneous colorectal cancer development, the adenoma-carcinoma sequence is a commonly used model. During the first stage, the development of an early adenoma from normal epithelium takes place. In 70 to 80% of all cases, this stage is caused by a somatic *APC* mutation in the glandular epithelium leading to the formation of adenomatous polyps.[5] These polyps are dysplastic and characterized by an altered differentiation of the epithelial cells. A germline mutation of *APC* is also responsible for an increased number of polyps (> 100) in the Familial Adenomatous Polyposis syndrome (FAP).[6,7] Due to a loss-of-function mutation, APC is not able to mediate the degradation of β-catenin in the absence of Wnt-ligands in the Wnt-signaling pathway.[6] Being a Wnt-signal dependent transcription factor, β-catenin induces, for example, the transcription of T-cell factor-regulated target genes (TCF).

The next stages in the adenoma-carcinoma sequence are the development of late adenomas and carcinomas caused by mutations in additional genes having key roles in tumor initiation, progression and maintenance. An increased mutation rate is caused e.g. by mutations of DNA mismatch repair genes (MMR-genes) like *MLH1*[8,9] and *MSH2*.[10,11] One consequence of this is the microsatellite instability phenotype (MSI), characterized by alterations in microsatellites within the genome.

Further examples of key alterations are mutations in spindle formation associated genes just as *Mad2*, *BubR1*, *Bub3* and *CENPE* leading to losses or gains of distinct chromosomal regions or complete chromosomes. The resulting phenotype is called "chromosomal instability phenotype" (CIN).[12] By these mechanisms, a huge number of different somatic mutations may evolve during cancer progression. In one case, Wood *et al.* identified more than 80 different somatic mutations within one primary colorectal cancer.[13] However, only few of them can be found in large proportions of different primary tumors. One of the most frequently altered genes is *TP53* encoding the tumor suppressor protein p53. Due to CIN leading to alterations at chromosome 17, Loss of Heterogeneity

(LOH) of *TP53* can be observed in 70 % of all CRC cases.[7] In the remaining allele, a somatic mutation can cause a complete loss of p53 function.[14] Because p53 is a key transcriptional regulator of genes associated with cell cycle checkpoints and apoptosis, loss of p53 leads to continued growth and reduced apoptosis after cellular stress like hypoxia or DNA strand breaks. Thus, the mutation rate increases and angiogenesis is favored,[15] circumventing hypoxia and leading to metastases.

Deregulation of components of the MAPK-pathway, responsible for cell proliferation, differentiation and inhibition of apoptosis, and the PI3K-pathway, also responsible for inhibition of apoptosis, are two other prominent examples of alterations occurring in cancer tissues.[16,17] Due to their importance, both pathways and their influence on colorectal cancer will be presented in detail in section 1.2. Furthermore, epigenetic alterations, comprising changes in chromatin structure, DNA methylation and miRNA-abundances, are crucial factors for the development of colorectal cancer. In section 1.4, the influence of these factors on gene expression and on the development of cancers, especially colorectal cancer, will be described.

1.2 The MAP-kinase and PI3-kinase pathway

1.2.1 The EGF-receptor activates MAPK- and PI3K-signaling

The epidermal growth factor receptor (EGFR) is a 170 kDA transmembrane glycoprotein.[18] It consists of an extracellular domain, a hydrophobic transmembrane domain and an intracellular domain. The extracellular domain is organized in 4 subdomains responsible for ligand binding. The intracellular domain includes a juxtamembrane domain, a tyrosin kinase domain and a long carboxyterminal region. This region encompasses autophosphorylation sites, which are binding sites for SH2-domains and PTB motifs. EGFR, also known as ErbB1, is one of the four members of the ErbB family of type I receptor tyrosin kinases. While ErbB2 and ErbB4 have similar structures to ErbB1 and are highly homologous in their tyrosin kinase domains, ErbB3 lacks this domain. In contrast, the extracellular regions responsible for ligand binding are less conserved between the family members. Therefore, each receptor monomer can bind different ligands called EGF-like growth factors. At least 12 different EGF-like growth factors were identified (table 1), of which four are able to bind EGFR, three are able to bind EGFR and ErbB4 and four are able to bind ErbB3 and ErbB4. The receptor bound by Cripto-1 is still unknown. Interestingly, ErbB2 is not able to bind any ligand, although it is the most potent dimerization partner for the other ErbB-receptor monomers.

When binding a ligand, EGFR either forms a homodimer with a second EGFR-monomer or a heterodimer with one of the other monomers. The composition of the dimer determines the effect of the ligand on the cell. If, for example, ErbB2 is present in the receptor dimer, ligand dissociation slows down leading to a prolonged signal to the cell.[19] Other factors are for instance the ligands' or the receptor monomers' abundances.

Table 1: EGF-like growth factors and receptors: An X indicates which growth factor is able to bind which receptor monomer.

EGF-like growth factor	EGFR	ErbB2	ErbB3	ErbB4
EGF	x			
TGF-α	x			
AREG	x			
Epigen	x			
EREG	x			x
BTC	x			x
HB-EGF	x			x
Neuregulin 1 (NRG1)			x	x
NRG 2 alpha and beta			x	x
NRG 3			x	x
NRG 4			x	x
Cripto-1 (CR-1)	unknown	unknown	unknown	unknown

1.2.2 Signal transduction after receptor activation

After dimerization, tyrosine residues in the carboxyterminal region of the receptor are autophosphorylated and prepared for binding of SH2-domain containing downstream proteins such as GRB2. GRB2 has two SH3-domnains being binding sites for the guanine nucleotide exchange factor SOS. By binding, SOS is activated and removes GDP from Ras (KRAS, NRAS, HRAS), enabling it to get activated by binding GTP. Active Ras binds and activates downstream targets like Raf-kinase (BRAF, CRAF, ARAF), which then phosphorylates and activates MEK (MEK1 and MEK2). Active MEK phosphorylates and activates MAPK (ERK1, ERK2) which can move into the nucleus and activate downstream factors like MYC (C-myc). MYC is mainly involved in regulatory processes of gene transcription responsible for cell proliferation.
Another mechanism of active receptor tyrosine kinases is to trigger the PI3K-pathway via binding of the adapter protein IRS. Active IRS binds PI3-kinase which phosphorylates PIP_2 to PIP_3. PIP_3 is able to activate AKT which then, for instance, inhibits pro apoptotic proteins like BAD.

1.2.3 Alterations of the MAPK- and PI3K-pathway in cancer development

The MAPK- and PI3K-pathway are two of the most important pathways in colorectal cancer, because they regulate proliferation, differentiation and apoptosis. Mutations of genes within these pathways lead to permanent activation of proteins causing growth advantages compared to normal cells. RAS and RAF are components of the MAPK-pathway whose genes are somatically mutated in a significant subset of all colorectal cancer cases. *KRAS* is reported to be mutated in about 50 % whereas *BRAF* is reported to be mutated in about 5 % to 15 % of all cases.[20,21] Furthermore, somatic mutation of *PI3K* can cause permanent activation of the PI3K-pathway. This mutation is present in

about 10 % to 15 % of all colorectal cancer cases.[20,21]

Permanent activation of both pathways is also caused by high abundances of the EGF-receptor and its ligands. EGFR is highly expressed in about 65 % to 70 % of all colorectal cancer cases.[22] Also, its ligands Amphiregulin and Epiregulin are reported to be overexpressed in colorectal cancer and other cancer types.[23]

1.3 Amphiregulin and Epiregulin are ligands of the EGFR

1.3.1 *AREG* and *EREG* gene properties

Amphiregulin, first discovered by Shoyab *et al.*,[24] and Epiregulin, first discovered by Toyoda *et al.*[25] are EGF-like growth factors binding EGFR. Both genes are located on Chromosome 4q13.3. *Betacellulin* and *Epigen*, encoding other EGF-like growth factors, are located on this chromosomal band, too. *Amphiregulin* consists of 6 exons, leading to two major splice variants. Whereas the most abundant AREG precursor protein, being 252 aminoacids long, is made from exons 1,2,3,4 and 5, the alternative precursor protein, being 274 aminoacids long, is made from exons 1,2,3,4 and 6. *Epiregulin* consists of 5 exons leading to a 169 aminoacid precursor protein. Both, AREG and EREG are transmembrane proteins, harboring an EGF-like domain containing three disulfite bonds, necessary for EGFR-binding. In contrast to AREG, EREG is also able to bind ErbB4 (see table 1). The precursor AREG protein (pro-AREG) contains several cleavage sites and glycosylation motifs. These motifs lead to different mature AREG proteins and influence AREG's biological activity in different cell types.[26] Primarily, ADAM-17, a metalloprotease localized on the cell surface, cleaves pro-AREG and releases a 84 aminoacid soluble biological active form into the extracellular space (shedding).[27] After shedding, the remainder of AREG protein can translocate into the nucleus and regulate global transcription.[28] Similar to AREG, the EREG precursor is also shedded by ADAM-17 to release a 46 aminoacid functional active form into the extracellular space,[29] enabling interaction with EGFR or ErbB4.

1.3.2 *AREG* and *EREG* gene function

Being EGFR or ErbB4 activating growth factors, AREG and EREG mediate proliferation and differentiation. When shedded by ADAM-17, they can function as autocrine or paracrine, if not shedded, as juxtacrine signal transducers.[30] By activating proliferation after estrogen and progesterone stimulation, AREG plays a central role in the development of the mammary glands during puberty and pregnancy.[31,32] AREG also mediates endometrium proliferation and preparation of embryo implantation into the uterus.[33,34] During oocyte maturation, *AREG* and *EREG* gene expression is increased.[35] AREG is involved in branching and tubulogenesis in several tissues, too, such as prostate,[36] kidney[37] and lungs.[38] Furthermore, *AREG* is regulated by gonadotropic hormones during spermatogenesis.[39] Finally, AREG is important in bone formation[40,41] and crucial in

neuronal development. It was identified to act as a growth factor in mouse adult neural stem cells.[42,43] In addition to the previously mentioned function in oocyte maturation, EREG was found to contribute to corneal epithelial wound healing[44] and wound healing in gingival cells.[45] *EREG* expression is also increased during vascular smooth muscle cell de-differentiation, necessary for vascular remodeling.[46]

1.3.3 *AREG* and *EREG* in malignant tissues

Besides their functions in tissue development, *AREG* and *EREG* were also identified to be expressed in a variety of different cancer types. *AREG* over-expression was, for example, observed in breast,[47,48] lung,[49,50] liver,[51,52] prostate,[53] pancreatic,[54,55] and colon cancer.[56,57,58] *EREG* over-expression was identified in breast,[59] lung,[60] colon[61] and other cancer types, too.[62,63] Although EREG was also reported to inhibit tumor derived epithelial cell lines,[25] the main function of both proteins is the stimulation of proliferation by activating EGFR. Understanding AREG and EREG regulation might help to identify mechanisms of tumor growth. To do so, addressing the epigenetic features of *AREG* and *EREG* genes might be promising, since epigenetics is a major driver for regulatory mechanisms of gene expression.

1.4 DNA methylation and histone modifications are epigenetic regulatory mechanisms

In 1939, Conrad Hal Waddington published his book "An Introduction to Modern Genetics" in which he first introduced and defined the phrase "Epigenetics".[64,65] Thus, epigenetics is almost 15 years older than Watson's discovery of the DNA-doublehelix.[66] Waddigton created the phrase "Epigenetics" as a name for a model to describe how genes might interact with their surroundings to produce a phenotype. Over the decades, the definition of "Epigenetics" developed and changed. In one example, epigenetics was described by Russo *et al.* as "The study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA-sequences."[67] Another way of describing epigenetics was made by Adrian Bird. He defined epigenetics as "The structural adaption of chromosomal regions so as to register, signal or perpetuate altered activity states".[68] The best understood mechanisms responsible for structural adaption are DNA methylation and histone modifications, both altering gene expression during development and cancer progression.[69] Of course, other epigenetic mechanisms are known which also influence gene-expression such as miRNA mediated gene silencing. However, due to their importance in this work, the explanations in the next sections focus on DNA methylation and histone modifications.

1.4.1 DNA methylation and histone modifications

DNA methylation was first discovered by Rollin Hotchkiss in calf thymus DNA in 1948.[70] Typically, it occurs at CG dinucleotides (CpGs) in the human genome. The enzymes responsible for DNA methylation are DNA methyltransferases (DNMTs), first discovered in 1975.[71] DNMT1,[72] responsible for maintenance of methylation during cell cycle, and DNMT3a and 3b[73],[74] responsible for *de novo* methylation, are the most important DNMTs in mammals. Genomic regions which can be methylated are e.g. located within the genes' promoter regions. There, CpG-rich regions, so called CpG islands, are regulatory sites capable of inhibiting gene transcription when methylated.[75] CpG-islands are characterized by a GC-content greater than 50 % with an observed to expected CpG ratio of greater than 60 % in a minimum of 200 bp region.[76] Other methylation-dependent events are genomic imprinting[77] or X-chromosome inactivation[78] causing chromosomal closure. However, the majority of methylation within the human genome was not discovered in CpG-islands, but in CpG-positions within CpG-poor regions (1 CpG per 100 bp).[79,80] In a genome wide approach, Brenet *et al.* recently discovered that first exon methylation is linked to transcriptional silencing.[79] Unfortunately, they were not able to present a model for the mechanism of intragene methylation-mediated gene expression, yet.

An additional means of DNA modification was discovered recently, called CpG hydroxymethylation. It is caused by oxidation of methylated CpGs by TET1.[81] Although little is known about CpG hydroxymethylation, it might contribute to gene regulation and will influence epigenetic research in the future.

Other important epigenetic features are histone modifications. Due to their capability to determine if DNA is accessible for transcription or not, they have a major impact on gene expression, too. First described by Albrecht Kossel in 1884,[82] histones are a protein class, containing histone H1 and the core histones H2A, H2B, H3, and H4. A nucleosome is formed by an octamer of the core histones and 146 basepairs of DNA.[83] Together with H1, the nucleosomes form the chromatin structure where DNA is condensed up to 10000-fold.[84] However, to allow processes like transcription or replication, this structure needs to change dynamically from a condensed state to locally open states. These dynamic changes are affected by three covalent histone modifications on the aminoterminal ends of the core histones: acetylation, phosphorylation and methylation. Besides some exceptions in yeast and drosophila,[85,86] histone acetylation is mainly linked to transriptional activation.[87] One important acetylation site is Lysine 14 at the aminoterminal end of histone 3 (H3K14).[88] By introducing or removing acetylgroups, histone acetyltransferases (HATs)[89] and histone deacetylases (HDACs)[90] change the affinity to DNA and adjacent nucleosomes.

Histone phosphorylation is also associated with transcriptional activation.[91] However, the mechanism is not well understood, yet. Similar to acetylation, phosphorylation might reduce the affinity to DNA and nucleosomes by introducing a negative charge. It

was also shown that phosphorylation stimulates HAT activity.[92]

Histone methylation can be associated either with transcriptional silencing (H3K9 di- or trimethylation)[93] or activation (H3K9 monomethylation[93] or H3K4 trimethylation[94]). One effect of methylated histones is, for instance, the binding of heterochromatin specific proteins like HP1 in silencing mechanisms.[95]

1.4.2 Epigenetic regulatory mechanisms

DNA methylation-dependent gene regulation is mainly mediated via two mechanisms. The first mechanism involves DNA methylation-dependent transcription factors including the most prominent ones, AP-2[96] and Sp1.[97] Both do not bind when the binding sites are methylated, leading to a reduction of gene expression.

The second mechanism comprises the connection between DNA methylation and chromatin structure, established by proteins binding to methylated DNA. One of the proteins is MeCP2[98] which has a methylation binding domain (MBD-domain)[99] and a transcriptional repression domain (TRD-domain). Thus, the protein is able to recruit a corepressor complex, consisting of mSin3A and HDACs, to methylated DNA. By deacetylation the chromatin changes to the condensed state leading to transcriptional silencing.[100] Besides of MeCP2, other MBD-containing proteins like MBD2 are able to recruit HDACs to methylated DNA.[101] The functional interaction between DNA methylation and histone modifications were also verified by studies showing that HDAC inhibitors like Trichostatin A are able to relief MBD-containing protein mediated transcriptional silencing.[102,101,103] Opposing this mechanism, chromatin structure can also influence DNA methylation. By trimethylating H3K27, polycomb group proteins, such as EZH2, induce transcriptional silent chromatin.[104] Furthermore, they are able to recruit DNMTs to the transcriptional silent sites[105] leading to methylated DNA.

In contrast to the MBD-containing proteins, the zinc finger containing proteins ZBTB33 (Kaiso)[106] and ZBTB4 are two members of another group of transcription factors. Their zinc finger motifs show higher affinity to methylated DNA sequences than to unmethylated DNA *in vitro* and thereby might repress transcription.[107] Additionally, ZBTB33 might be able to suppress gene expression also methylation-independently through extra zinc finger binding motifs, called BTB/POZ-domains.[106] The described functions of ZBTB33 include, for instance, the recruitment of the N-CoR repressor complex, a protein complex which promotes histone deacetylation leading to silent chromatin.[108] However, in a recent publication it was shown that the methylation-dependent binding might only play a minor role in the function of ZBTB33.[109] By analyzing ENCODE-data[110] the authors rather identified an association of ZBTB33-binding and actively expressed genes.

CTCF is also a zinc finger containing DNA-binding factor, which can bind to CpG-containing sites and to sites without CpGs, too.[111] Contrary to ZBTB33 and ZBTB4, methylation of the CpG-containing binding sites reduces binding affinity.[112] CTCF can

either function as a transcriptional repressor[113] or as a transcriptional activator.[114] The most important function of CTCF is to be an insulator protein being necessary to block promoters from the influence of functionally-independent enhancers.[115] CTCFL (BORIS), a paraloque of CTCF, is able to bind the same DNA motifs as CTCF.[116] Detected in several cancer cells,[117] CTCFL is thought to interfere with CTCF-binding and to function as an antagonist to CTCF.[118]

1.4.3 The impact of epigenetic mechanisms in cancer development

Epigenetic mechanisms play a major role in the formation and maintenance of all types of cancers. In 1983, it was detected that tumor tissues have a globally reduced methylation content in comparison to corresponding normal tissues.[119,120] The results were later verified by high throughput DNA methylation analysis techniques like microarrays.[121] This feature is called hypomethylation and is mainly found in gene poor areas[122] but sometimes, it can also occur at CpG-islands in promoters[123] when growth-related genes are affected. The *PAX2* gene promoter e.g. was found to be hypomethylated in endometrial cancers, but not in normal endometrial tissues.[124] Besides growth activation, hypomethylation also plays a role in chromosomal instability[125,126] and chromosomal rearrangement,[127] which is for example observable in the reactivation of transposons.[128] A further consequence of this, is loss of imprinting which can be seen, for instance, for the insulin-like growth factor 2 in colorectal cancers (*IGF2*).[129,130] During cancer development DNA hypomethylation increases.[131] In contrast to hypomethylation, hypermethylation mainly occurs in gene promoters. In different cancer types, the promoters of the genes *BRCA1*,[132] *p16*,[133,134] *E-cadherine*[135] and *VHL*[136] can be methylated, leading to their down-regulation. The hypermethylation pattern is specific for the cancer type and increases during cancer development.[137,138]

For colorectal cancer, a particular hypermethylation pattern is described as CpG-island methylator phenotype (CIMP). The CIMP-phenotype, first described by Toyota *et al.* in 1999, is characterized by hypermethylation of several promoter CpG-islands associated with inactivation of tumor suppressor genes.[139] Toyota *et al.* suggested a list of 30 different CpG-islands to distinguish tumors in CIMP positive or CIMP negative tumors. Since its discovery, the CIMP-phenotype was associated with epidemiological features, like age, gender or location of the tumor, as well as genetic features, like MSI, *KRAS* and *BRAF* mutations.[140,141] For example, when the *hMLH1*-promoter is among the methylated loci, CIMP positive tumors mainly have an MSI phenotype. But, in contrast to MSI or other well known phenotypic subtypes of cancers, CIMP still is under controversial discussion and not accepted by the entire scientific society.[142,143] The main reason is that the sites determining CIMP phenotype were not standardized, yet. For the 30 different CpG-islands of the analyzed CIMP positive tumors a variety of different methylation patterns is observable.[139] As a consequence, each individual tumor might have a different expression pattern of the 30 genes whose promoters are tested. Also, the

mechanism for CIMP tumor development is still unknown. Therefore, there is no clear argument supporting the relevance of the tested CpG-islands for predicting the CIMP status.

A further example of promoter methylation leading to silenced gene expression is the retinoblastoma gene (*RB*).[144,145] Silenced RB cannot contribute to a protein complex containing the chromatin remodeling proteins SWI/SNF,[146] HDACs,[147] polycomb class epigenetic silencing proteins[148] and DNMT1.[149] The complex is necessary to silence RB targets as e.g. the cell cycle activator gene *Cyclin E*[150] by changing the chromatin structure. This is an important example illustrating that DNA methylation is tightly linked to chromatin structure in cancer cells, too. The findings are further supported by changes in histone patterns like global loss of H4K16 acetylation or H4K20 trimethylation. These modifications were identified as common characteristics of cancer cells.[151] Additionally, the expression of histone modifying enzymes differs between healthy tissues and cancer tissues and also between the cancer types.[152] As these examples illustrate, epigenetic features play an important role in cancer research. Since these features are also connected to cancer development or the response of cancer tissues to medication, epigenetics also raises importance in clinical setting as prognostic, diagnostic and predictive markers.

1.5 Cancer treatment

1.5.1 Standard treatment options for cancer patients

Cancer can be treated by various therapeutic approaches. The most common approach is a surgery where the tumor and close lymph nodes are removed.[153] However, if the tumor is located in a non-accessible region or if there is a high estimated risk that the tumor relapses, further treatment options are applied. Radiation therapy is a means of killing tumor cells by electromagnetic or particle radiation.[154] The advantage of this is that radiation can be applied directly to the tumor with optimized intensities for the individual patient. The disadvantage is that surrounding healthy tissues might be harmed by the radiation, too, leading to a variety of side effects like hair loss, damaged organs or recurring cancer. A further therapy approach is chemotherapy. Chemotheapy is used to kill fast proliferating cells, like tumor cells, by compounds influencing cell proliferation.[155] One example is the FOLFIRI therapy in colorectal cancers which combines folinic acid, 5-fluorouracil and irinotecan hydrochloride.[156] Folinic acid interferes with nucleotide synthesis processes. 5-fluorouracil is a substitute of Cytosine and Thymine during DNA polymeration processes. Irinotecan hydrochloride is a topoisomerase inhibitor. Due to their impact on DNA replication and cell cycle processes, these compounds do not only kill tumor cells but also other fast proliferating cells like blood cells. This leads to side effects which are for example fatigue, digest problems or hair loss. To circumvent these problems, new therapeutics were tested which inhibit cell proliferation by targeting a tumor specific feature. These targeted therapeutics might be able to only affect the

tumor and not to influence normal tissue cells.

1.5.2 EGFR-targeted therapies

Over-expression of the EGF-receptor is a feature, widely spread among different tumor types. EGFR-inhibition will immediately lead to inhibition of cell growth and differentiation. Therefore, several compounds targeting EGFR were developed.

Cetuximab Also called Erbitux®, the most important therapeutic to target EGFR is Cetuximab. Cetuximab is a chimeric human/mouse monoclonal antibody. The mouse precursor antibody mAB225 was developed by Gill *et al.* in 1984.[157] They already discovered that mAB225 is able to inhibit EGF-binding and to block EGFR autophosphorylation in A431 cells, an epidermoid carcinoma cell line with EGFR over-expression. In subsequent experiments, it was investigated that mAB225 also inhibits breast cancer,[158,159] colon cancer,[160,161] renal cancer[162] and prostatic cancer cell lines.[163,164] Experiments in xenografts[165,166] showed similar results and paved the way for the first phase I clinical trial. Patients with squamous cell carcinomas of the lung were treated with ^{111}In-labeled mAB225 to test for toxicity and to visualize the tumor.[167,168] The great advantages of the antibody therapy was that no toxicity was observed when patients were treated with high doses and that the antibody targeted the lung tumor cells directly. This was observed by radiolabel-dependent visualization three to five days after injection.

In 1993, the mAB225 was chimerized by Naramura *et al.*[169] to avoid human anti-mouse antibody response (HAMA).[170] Similar to mAB225, the mouse-human chimeric C225, now called Cetuximab, also inhibited cell growth in several cancer cell lines and xenografts.[171,172,173,174] As a consequence, numerous phase II and phase III clinical trials were performed. In most of the cases Cetuximab was combined with other drugs or treatment options like radio- or chemotherapy. In a phase III clinical trial from 2005, it was seen that initial treatment of metastatic colorectal cancer with Cetuximab plus chemotherapy (FOLFIRI) reduced the risk of disease progression by 15 % in comparison to treatment with FOLFIRI alone (CRYSTAL-study).[175] In this study, it was also seen that *KRAS*-wildtype patients responded significantly better to chemotherapy plus Cetuximab than *KRAS*-mutant patients. However, also among the *KRAS*-mutant patients a subset of patients responded better to Cetuximab plus FOLFIRI (odds ratio (95 % CI) = 0.80 (0.44 - 1.45), odds ratio > 1: benefit from Cetuximab plus FOLFIRI). Finally, as a result of this study, Cetuximab was approved in the USA and in Europe for the treatment of metastatic colorectal cancers e.g. in combination with FOLFIRI when the tumor highly expresses *EGFR*. Yet, this treatment option is only valid for *KRAS*-wildtype patients.

Besides Cetuximab, other antibodies were developed to interfere with EGFR-function in colorectal cancer. One example is Panitumumab (Vectibix®) which was

also approved for monotherapy and in combination with chemotherapeutics. However, small molecules targeting the intracellular tyrosine kinase domain of the EGFR are also promising. The most widely used ones are Erlotinib and Gefitinib.

Erlotinib and Gefitinib Erlotinib was shown to induce cell cycle arrest and apoptosis in cancer cell lines.[176] Also, Gefitinib was shown to inhibit the MAPK-pathway in high *EGFR* expressing cancer cells.[177] Both compounds were subjected to several clinical trials. In contrast to Cetuximab, Erlotinib as well as Gefitinib failed to increase response in combination therapies with chemotherapeutics.[178,179,180,181] However, as a monotherapy, Erlotinib succeeded to increase the life-span of NSCLC-patients.[182] Gefitinib also led to a better outcome, but only in patients harboring *EGFR* mutations.[183] The SATURN-trail confirmed an increased effect for patients with *EGFR* mutations, too.[184] As a consequence of these studies, FDA and EMA approved Erlotinib for the treatment of non-small cell lung cancer and pancreatic cancer. Gefitinib was also approved for the treatment of non-small-cell lung cancer, if the tumor comprises an activating *EGFR* mutation. However, none of the compounds were approved for treatment of colorectal cancer until now. But, phase II clinical trials are in progress.[185,186,187,188]

Targeted therapies gain importance in clinical treatment procedures. Patients are stratified with biomarkers like *KRAS*- or *EGFR* mutations to increase the response rates of such therapies. However, other biomarkers will be necessary to further optimize the stratification procedure for example to exclude patients who will not respond, although they are *KRAS*-wildtype.

1.5.3 *AREG* and *EREG* expression are predictive markers for EGFR-targeted therapies

Due to their influence on EGFR-activation leading to activated PI3K- and MAPK-pathways, AREG and EREG might interfere with inhibitory mechanisms of EGFR inhibitors. Because of this, they were considered as potential predictive markers to evaluate the outcome of EGFR-targeted therapies. Several studies were e.g. performed in non-small cell lung cancers. On the one hand, Ishikawa *et al.* and Masago *et al.* showed that AREG expression correlates with a poor response to Gefitinib in NSCLC-patients.[189,190] On the other hand, Yonesaka *et al.* found out that Gefitinib as well as Cetuximab led to a higher growth inhibition in high AREG expressing NSCLC-cell lines compared to low AREG expressing NSCLC-cell lines.[191] They verified the cell line data by immunohistochemistry analyses in patient-derived NSCLC-tumors. High AREG-staining indicated a stable disease after Erlotinib or Gefitinib treatment, whereas low AREG expressing patients mainly showed progressive disease. Vollebergh *et al.* also verified these results by testing the sera of a large patient cohort. High AREG levels were associated with a significantly decreased risk of death after Erlotinib or Gefitinib treatment.[192] Interestingly, in the same study, low TGF-α levels correlated with a better prognosis than high

TGF-α levels.

In metastatic colorectal cancers, AREG and additionally EREG were also identified as predictive markers for the outcome of EGFR-targeted therapies. In a 110-patient-study, Khambata-Ford *et al.* observed that high *AREG* mRNA and high *EREG* mRNA expression levels were mainly detected in patient biopsies which responded to Cetuximab therapy. They also showed that high *AREG* and *EREG* mRNA expression correlates with a longer progression-free survival time. Finally, they verified previous data illustrating that *KRAS*-wildtype patients responded better to Cetuximab than *KRAS*-mutant patients.[193] However, in their study, the mutation status of the patients did not influence the potential of AREG and EREG as predictive markers. In contrast to these results, Jacobs *et al.* observed that the *KRAS* mutation status is indeed a criterion for the potential role of *AREG* and *EREG* as predictive markers. In their study, including 220 patients, high *AREG* as well as high *EREG* mRNA expression correlated with the response towards a therapy including Cetuximab and Irinotecan in *KRAS*-wildtype patients.[194] In a study, containing 226 patients, Pentheroudakis *et al.* verified the findings of Jacobs *et al.* for AREG.[195] However, in contrast to the study performed by Jacobs *et al.*, they additionally observed that *EREG* mRNA expression correlates with the response towards Cetuximab irrespective of the *KRAS* mutation status.[195] That EREG can also be a prognostic marker for overall survival of colorectal cancer patients who never received EGFR-targeted therapy, was shown in another trial by Kuramochi *et al.*[52] Similar to Pentheroudakis *et al.*, high *EREG* mRNA expression correlated with a higher overall survival rate in mutant *KRAS*-patients. Interestingly, in *KRAS*-wildtype patients *EREG* mRNA expression correlated with a lower overall survival rate. A study performed by Yoshida *et al.* verified the potential of AREG and EREG, but also the potential of other EGFR-ligands to predict the response towards EGFR-targeted therapies.[196] The results were obtained by immunohistochemistry experiments. Therefore, it would be easy to include AREG and EREG evaluation in routine diagnostic procedures without the need to purify RNA. However, although the studies mentioned above mainly point towards a positive therapy response of colorectal cancers with high AREG and EREG levels when treated with EGFR-targeted therapeutics, it is still under investigation how other tumor characteristics, like mutations, might influence these findings. Therefore, the identification of the biological background of the tumor in combination with the AREG and EREG expression and their regulation mechanisms is very important.

1.5.4 Epigenetic markers for diagnosis and cancer treatment

As already mentioned, the CpG island promoter methylation patterns varies between healthy and cancer tissues and also among different cancer types.[137] Therefore, the identification of specific methylation patterns could serve as a good diagnostic tool for cancer characterization. One example is the determination of the CIMP-status in colorectal cancer patients described in section 1.4.3.

Nevertheless, the methylation status of single loci might also contribute to cancer diagnosis. The methylation status of the Glutathione S- transferase P1 promoter (*GST1P*), for example, can be used to distinguish malignant from benign tissues. In prostate cancer patients, the *GST1P*-promoter is mainly hypermethylated[197] but in benign hyperplastic prostate tissues it is often unmethylated.[198] As a second example, promoter methylation of the *BRCA1* gene occurs in an early stage of hereditary breast cancer development.[199] Therefore, it might contribute to the diagnosis of breast cancer onset in high-risk patients. As a further advantage, several studies showed that methylation analyses might be easily applied on biopsies or biological fluids.[200,201,202]

In prognostic procedures, epigenetic features can also be used to predict the potential course of disease. In head and neck cancer, methylation of the *p16* promoter correlates either with a decreased survival in *TP53*-wildtype patients or with a better survival in *TP53*-mutant patients.[203] Also, in colorectal cancers, *p16* promoter methylation correlates with a poor prognosis.[204] Another example is the DNA repair gene O6-methylguanine-DNA methyltransferase (*MGMT*). A correlation was found between promoter hypermethylation of this gene and a poor prognosis of patients with brain cancers.[205] Interestingly, in B-cell lymphomas, hypermethylation of the same gene correlates with an increased overall survival.[206]

MGMT promoter methylation might also play a role in predicting the outcome of cancer treatment procedures. In gliomas, it corresponds with the clinical response to alkylating agents like carmustine[207] or temozolomide.[208] Other examples are hypermethylation of the *hMLH1* promoter, which correlates with resistance to cisplatin-based chemotherapy in ovarian cancer[209] and hypermethylation of the reduced folate carrier gene (*RFC*), which is associated with a diminished response to methotrexate-based chemotherapy in lymphomas of the central nervous system.[210]

The impact of epigenetic features as biomarkers during clinical evaluation processes raised in the last years and might become as important as mutation patterns. However, the difference between mutations and epigenetic features is that the latter can be altered by epigenetic compounds such as DNA methyltransferase inhibitors or HDACis. Thus, epigentically silenced genes could be reactivated, which is not possible for mutated genes. Therefore, it is not surprising that epigenetic compounds are also discussed as potential drugs in cancer treatment. DNA methyltransferase inhibitors, such as 5-Azacytidine (Vidaza) or 5-Aza 2'deoxycytidine (DAC), have already been approved for the treatment of myelodysplastic syndrome and leukemia.[211] Zebularine has entered clinical trials, too.[212] Also HDACis were already approved for cancer treatment. Vorinostat and Depsipeptide are for example used for treating cutaneous T-cell lymphoma.[213,214] Other HDACis are subjected to phase I[215,216,217,218] and phase II clinical trials.[219,220,221] But in monotherapy, these compounds mainly show little or no clinical activity.[222,223,224] However, preclinical data suggests that combination therapies using HDACis together with other inhibitors might increase response.[225,226,227,228,229]

1.6 Future prospects

Cancer is a heterogeneous disease having many molecular characteristics. To cope with this heterogeneity, cancer is normally treated with chemotherapeutics or by radiotherapy in order to destroy unspecifically most of the tumor cells. The introduction of targeted therapies, such as Cetuximab in colorectal cancers, led to a more specific way of cancer treatment. However, patients which might benefit from such a treatment option must fulfill several molecular requirements, such as the *KRAS*-wildtype status or *EGFR*-expression. Nevertheless, also a subset of patients, which do not fulfill the requirements might benefit from that treatment. On the contrary, a subset of patients treated, does not respond, but might respond to other yet unknown treatment options. To circumvent this dilemma, several research projects were launched aiming for deeper insights into the molecular mechanisms of cancer cells. The international Cancer Genome Consortium e.g. was founded to analyze 50 different cancer genomes for their genomic, transcriptomic and epigenomic changes in comparison to normal tissues. Another project is the ColoNET consortium, founded to create an *in silico* model of the most important signaling pathways involved in processes within colorectal cancer cells. These pathways might contribute to identify new potential biomarkers in tumor therapy. Data was first collected by analyzing well characterized colorectal cancer cells *in vitro* and *in vivo*. In subsequent experiments, these findings should be approved in tumor tissues. Two of the potential biomarkers identified were AREG and EREG whose expression correlates with the outcome of EGFR-targeted therapies.[193,194]

1.7 Aim of the work

Within this work, the epigenetic regulation of the EGFR-ligands *Amphiregulin* and *Epiregulin* was analyzed. Different colorectal cancer cell lines were tested for their *AREG* and *EREG* mRNA and protein expression and the alteration of expression after treatment with DNMT inhibitors and HDAC inhibitors. These results led to the following questions:

1) Where are the epigenetic regulatory regions within the *AREG* and *EREG* genes?
2) Do these regulatory regions have promoter function?
3) What is the mechanism of epigenetic regulation?
4) Are the epigenetically regulated regions useful biomarkers for EGFR-targeted therapies?
5) Is it possible to increase therapy sensitivity by increasing *AREG* or *EREG* expression, either by epigenetic alteration or by over-expression experiments?
6) Can the *in vitro* findings be adopted to *in vivo* systems?

2 Material and Methods

2.1 Materials

2.1.1 Consumables

Table 2: Consumables

Material	**Company**
0.2 µm syringe filters	Whatman
0.5 µl PCR tubes	Applied Biosystems
10 µl, 100 µl, 300 µl, 1250 µl filter tips	Sarstedt
10 µl, 200 µl, 1000 µl pipet tips	Eppendorf
1250 µl Bulk Tips	Matrix
2 ml, 5 ml, 10 ml 25 ml pipet tips	BD Falcon
6-well plates	BD Falcon
24-well plates	BD Falcon
96-well-plates	BD Falcon
96-well PCR Platte, farblos "Fast" Typ	Biozym
96-well-plate Nunc MaxiSorp® flat-bottom	Thermo Scientific
Adhesive Clear qPCR Seals	Biozym
Cell culture flasks 75 cm^2	BD Falcon
Cell scraper 25 cm	Sarstedt
Cryovials 2ml	Cellstar
polypropylene tubes 5 ml, 75 x 12 mm	Sarstedt
Illustra MicroSpin G-50 Columns	GE Healthcare
Lab glasware	Duran
MicroAmp® 8-Cap Strip	Applied Biosystems
MicroAmp® Clear Adhesive Film	Applied Biosystems
MicroAmp® Fast 8-Tube Strip 0.1 ml	Applied Biosystems
MicroAmp® Fast Optical 96-well Reaction Plate 0.1 ml	Applied Biosystems
Nylon transfer membranes Hybond N+	Amersham pharmacia
Petri dishes 10 cm	BD Falcon
Reaction tubes 1.5 ml, 2 ml	Eppendorf
Reaction tubes 15 ml, 50 ml	BD Falcon
Whatman 1MM papers	Whatman

2.1.2 Devices

Table 3: Devices

Devices	Company
8-Channel Electronic Pipette, 15 - 1250 µl	Matrix
8-Channel Multipipet 200 µl	Eppendorf
Agarose gel chamber	Bio-Rad
Agarose gel documentary system	Biostep
Balance BP 2100 S	Sartorius
Balance RC 210 P	Sartorius
BD FACSCalibur^{TM}	BD biosciences
Benchmark Plus Microplate Spectrophotometer	Bio-Rad
Bio-Plex 200	Biorad
Cell Counter TC10	Bio-Rad
Centrifuge Allegra^{TM} 6R	Beckman Coulter
Centrifuge Allegra^{TM} X15R	Beckman Coulter
Centrifuge 5415C	Eppendorf
Centrifuge 5417R	Eppendorf
Centrifuge 5424	Eppendorf
Centrifuge Avanti^{TM} J25	Beckman Coulter
Clean bench LaminAir® HBB2448	Heraeus instruments
Clean bench Class II	NUIARE
Clean bench HS12	Hera
Cryo 1 ℃ freezing container	Nalgene
Hybridisation oven 6V/12V	Unitherm
Hypercassette	Amersham
Incubator GFL 3031	GFL
Incubator Hera cell 240	Hera
Luminoscan RS	Labsystems
Lumat LB 9507	Berthold Technologies
MicroAmp® Adhesive Film Applicator	Applied Biosystems
Microscope Leica DMIL	Leica
Nanophotometer	IMPLEN
Neubauer Improved cell counting chamber	Carl Roth
PCR cycler T1 Thermal Cycler	Biometra
Ph-meter CG 840	Schott
Pipet Boy	Eppendorf
Pipets: 0.5 - 10 µl, 10 - 100 µl, 100 - 1000 µl	Eppendorf
Plate shaker IKA® MTS 2/4	IKA
Power Supply EPS 200	Pharmacia Biotech
Qualitron® DW-41 Microcentrifuge	Krackeler Scientific
Real-Time PCR system StepOne^{TM} Plus	Applied Biosystems
Sonicator Transsonic T310-H	Elma
Stirrer IKAMAG RET	IKA
Thermomixer 5436	Eppendorf
UV-crosslinker	Hoefer
Vortexer Reax 2000	Heidolph
Water bath GFL 1003	GFL

2.1.3 Chemicals and solutions

Table 4: Chemicals

Chemicals	**Company**
2' Propyl valeric acid (Valproat, VPA)	Sigma
3,3' ,5,5' Tetramethylbenzidine (TMB)	Sigma
3-(N-morpholino)propanesulfonic acid (MOPS)	Fisher Scientific
5-aza-2' deoxicytidine (Decitabine, DAC)	Sigma
α-D-Glucose	Serva
[α-^{32}P]-UTP	Hartmann Analytic
Agarose	Serva
Albumin bovine (BSA)	Serva
Bacto^{TM} Agar	BD
Bacto^{TM} Tryptone	BD
Bacto^{TM} Yeast Extract	BD
BD FACSClean^{TM}	BD
BD FACSFlow^{TM}	BD
Calcium chloride ($CaCl_2$)	Merck
Calcium phosphate ($CaPO_4$)	Merck
Cambinol	Sigma-Aldrich
Cetuximab	Charité pharmacy
Dimethyl sulfoxide (DMSO)	Sigma
Disodium hydrogen phosphate (Na_2HPO_2)	Merck
dNTPs	Roche
Ethylene diamine tetraacetic acid (EDTA)	Sigma
Ethylene glycol tetraacetic acid (EGTA)	Merck
Erlotinib	LC Laboratories
Ethanol	J.T.Baker
Ethidium bromide	Sigma
Fetal Calf Serum (FCS)	Biochrom AG
Formaldehyde (37 %)	J.T.Baker
G418	PAA
Gefitinib	LC Laboratories
Glacial acetic acid	Merck
Glycerin	Merck
Hydrogen peroxide (H_2O_2)	J.T.Baker
Isopropyl alcohol (IPA)	J.T.Baker
Magnesium chloride ($MgCl_2$)	Merck
Magnesium sulfate ($MgSO_4$)	Merck
Manganese(II)chloride ($MnCl_2$)	Sigma
Methanol	J.T.Baker
Piperazine-1,4-bis(2-ethanesulfonic acid) (PIPES)	Merck
Polyoxyethylene (20) sorbitan monolaurate (Tween®-20)	Serva
Potassium dihydrogen phosphate (KH_2PO_4)	Merck
Potassium acetate (CH_3COOK)	Merck

Chemicals	**Company**
Potassium chloride (KCl)	Merck
Propidium iodide (PI)	Fluka
Protamine sulfate	Sigma
Rubidium chloride (RbCl)	Sigma
Sodium acetate (CH_3COONa)	Merck
Sodium chloride (NaCl)	Merck
Sodium nitrate ($C_6H_5Na_3O_7$)	Merck
Suberoylanilide hydroxamic acid (SAHA)	Cayman
Sulfuric acid (H_2SO_4)	Roth
Trichostatin A (TSA)	Enzo Life Sciences
Tris-Base	Merck
Triton X-100	Sigma
Ultraglutamine	Lonza
Xylene	J.T. Baker

Table 5: Commercial solutions

Commercial solutions	**Company**
1 kb DNA ladder	NEB
2x TaqMan® Gene Expression Mastermix	Applied Biosystems
2x Power SyBrGreen® Mastermix	Applied Biosystems
2x GoTaq® qPCR Mastermix	Promega
100 bp DNA ladder	NEB
100 bp DNA ladder	PeqLabs
6 x DNA loading dye	Thermo Scientific
Ampicilin (50 mg/ml)	Sigma
Antibody Diluent	Zytomed
Bio-Plex® cell lysis buffer	Biorad
Bio-Plex® detection buffer	Biorad
Bio-Plex® resuspension buffer	Biorad
Bio-Plex® wash buffer	Biorad
cOmplete, Mini, +EDTA	Roche
DMEM	Lonza
IPTG (100 mM)	Sigma
Kanamycin (50 mg/ml)	Sigma
Lipofectamine^{TM} 2000	Invitrogen
Liquid DAB + Substrate Chromogen System	Dako
NaCl-solution	PeqLabs
Optimem	Gibco
Penicillin/Streptomycin (10 mg/ml)	Biochrom
Peroxidase-Blocking Solution, Dako Real^{TM}	Dako
Sodiumpyrovate solution NaPyr (100 mM)	Gibco
Trypsin/EDTA (0.05%/0.02%)	Biochrom
Ultraglutamine (200 mM)	Lonza
ULTRAhyb^{TM}-buffer	Ambion
Puromycin dihydrochloride	Sigma
Vitro-Clud®	R.Langenbrinck
VLE RPMI 1620	Biochrom AG
X-gal (40 mg/ml)	Roth
Zeocin (100 mg/ml)	Invitrogen

Table 6: Solutions

Solutions	**Receipt**
1x LB medium	Bacto^{TM} Tryptone (10 g), Bacto^{TM} Yeast Extract (5 g), NaCl (10 g), pH = 7.5, ad 1 l ddH_2O
1x lowsalt LB medium	Bacto^{TM} Tryptone (10 g), Bacto^{TM} Yeast Extract (5 g), NaCl (5 g), pH = 7.5, ad 1 l ddH_2O
1x LB agar or 1x lowsalt LB agar	ad 15 g Bacto^{TM} Agar per 1 l 1x LB or 1x lowsalt LB
10x MOPS	MOPS (0.2 M), CH_3COONa (80 mM), EDTA (10 mM)
10x Tris EDTA buffer (TE)	Tris-Base (0.1 M), EDTA (0.01 M), pH = 8.0
20x SSC	per 1 l: 175.3 g NaCl, 88.2 g $C_6H_5Na_3O_7$ (Sodium citrate), dissolve in RNase-free water
50x Tris-acetate EDTA buffer (TAE)	Tris-Base (242 g), Glacial acetic acid (57.1 ml), EDTA (18.6 g), pH = 7.5 - 7.8, ad 1 l ddH_2O
Citrate buffer (IHC)	10mM Citric Acid, 0.05 % Tween 20, pH 6.0
D10-medium	DMEM (500 ml), FCS (10 %), Ultraglutamine (1 %), Pen/Strep (0.1 mg/ml)
ELISA: Reagent Diluent	1% BSA in 1x PBS, pH = 7.4, 0.2 µm filtered
ELISA: Stop Solution	2N H_2SO_4
ELISA: Wash Solution	0.05% Tween®-20 in 1x PBS, pH = 7.4
FACS-dilution buffer	0.1 % Triton X-100, 0.5 % BSA in 1x PBS pH = 7.4
Glycerin stock solution	Glycerin (65 %), $MgSO_4$ (0.1 M), Tris-HCl (0.05 M), pH = 8
Homogenization buffer	PIPES (10 mM), NaCl (400 mM), EGTA (1 mM), $MgCl_2$ (3 mM), pH = 7.4
N1 buffer (Northern blot)	2x SSC, 0.1 % SDS ad H_2O
N2 buffer (Northern blot)	0.1x SSC, 0.1 % SDS ad H_2O
Phosphate buffered saline (PBS)	NaCl (137 mM), KCl (2.7 mM), Na_2HPO_4 (10 mM), KH_2PO_4 (1.8 mM) pH = 7.4

Solutions	Receipt
RNA-loading buffer	per 2 ml: 1310 µl RNAse-free water, 210 µl 10x MOPS, 350 µl formaldehyde, 130 µl 80 % glycerin, 2 µl Ethidium bromide (10 mg/ml), 0.03 % Bromphenol Blue
RPMI-medium	VLE RPMI 1640 (500 ml), FCS (10 %), Ultraglutamine (2 mM), Pen/Strep (0.1 mg/ml), NaPyr (1 mM)
SOC-medium	Bacto Tryptone (2 %), Bacto Yeast Extract (0.5 %), NaCl (8.6 mM), KCl (2.5 mM), Mg_2SO_4 (20 mM), α-D-Glucose (20 mM)
Tris-buffered saline (TBS)	Tris-base (50 mM), NaCl (150 mM), pH = 7.6
TFBI	CH_3COOK (30 mM, pH = 5.8), RbCl (100 mM), $CaCl_2$ (10 mM), $MnCl_2$ (50 mM), Glycerol (15 %)
TFBII	Mops (10 mM, pH = 7), RbCl (10 mM), $CaCl_2$ (75 mM), Glycerol (15 %)

2.1.4 Commercial kits

Table 7: ELISA kits and antibodies

Target	Antibody	Company
Human AREG polyclonal AB, goat IgG	AB-262-NA	RnD systems
Rabbit anti goat HRP-conjugated antibody	61-1620	Invitrogen
Target	**ELISA kits**	**Company**
AREG	Duoset® dy262	RnD systems
BTC	Duoset® dy261	RnD systems
EGF	Duoset® dy236	RnD systems
EREG	E91945Hu	USCN
TGF-α	Duoset® dy239	RnD systems

Table 8: Bio-Plex® assays

Protein	Phosphorylation site	Beadnumber	Company
AKT	Ser473	75	Biorad
ERK1/2	Thr202/Tyr204, Thr185,Tyr187	38	Biorad
MEK1	Ser217/Ser221	40	Biorad
EGFR	Tyr	20	Biorad

Table 9: TaqMan Gene Expression Assays: All assays are human specific unless otherwise indicated.

Gene	Dye	Assay ID	Company
AREG	VIC	Hs00155832_m1	Applied Biosystems
AREG(mouse)	FAM	Mm00437583_m1	Applied Biosystems
BTC	FAM	Hs01101204_m1	Applied Biosystems
CTCF	FAM	Hs00902008_m1	Applied Biosystems
CTCFL	FAM	Hs00966548_g1	Applied Biosystems
EREG	VIC	Hs00914313_m1	Applied Biosystems
EREG(mouse)	FAM	Mm00514794_m1	Applied Biosystems
UBE2D2	FAM	Hs00366152_m1	Applied Biosystems
ZBTB33	FAM	Hs00272725_s1	Applied Biosystems
ZBTB4	FAM	Hs00394164_m1	Applied Biosystems

Table 10: Other commercial kits

Other commercial kits	**Company**
T7 High Yield RNA Synthesis Kit	NEB
BCA protein assay kit	Thermo Scientific
CalPhos^{TM} kit	Clontech
Cell Proliferation Kit II (XTT)	Roche
Developer and Replenisher X-OMAT EXII	Kodak
Dual-Luciferase® Reporter (DLR^{TM}) Assay	Promega
Fixer and Replenisher RP X-OMAT LO	Kodak
Gateway® Technology with Clonase^{TM} II	Invitrogen
High Pure RNA isolation kit	Roche
Phusion site-directed Mutagenesis kit	Thermo Scientific
PolyATtract® mRNA Isolation System IV	Promega
Qiagen® Multiplex PCR Kit	Qiagen
Qiagen® EpiTect Bisulfite Kit	Qiagen
Qiagen® Plasmid Maxi kit	Qiagen
Qiagen® Plasmid Midi kit	Qiagen
Qiaprep Spin Miniprep Kit	Qiagen
QIAquick® Gel extraction kit	Qiagen
QIAquick® PCR purification kit	Qiagen
QIAshredder kit	Qiagen
RNase-free DNase Set	Qiagen
RNeasy® Mini Kit	Qiagen
Transcriptor High fidelity cDNA synthesis kit	Roche
Venor® GeM Mycoplasm Detection kit	Minerva Biolabs

2.1.5 Enzymes

Table 11: Commercial enzymes: Enzymes were purchased together with suitable buffers and substrates

Enzymes	Buffer and Substrates	Company
AmpliTaq® Gold	10x Buffer, dNTPs	Applied Biosystems
AvaI	NEB-buffer 4	NEB
BamHI	NEB-buffer 2,3,4	NEB
CIP	NEB-buffer 2,3,4	NEB
EcoRI	NEB-buffer 2	NEB
EcoRV	NEB-buffer 3	NEB
HindIII	Buffer R	Fermentas
DNA Pol. I lg Fragment (Klenow)	NEB-buffer 2	NEB
M.SssI	NEB-buffer 2, 32 mM SAM	NEB
NcoI	NEB-buffer 1,2,3,4	NEB
Pfu Ultra^{TM} DNA Pol.	10x buffer, 10 mM dNTPs,	Agilent Technologies
Phusion® HF DNA Pol.	5x HF buffer, 10 mM dNTPs	Thermo Scientific
T3-RNA-polymerase	RNAPol Reaction buffer	NEB
T4-ligase	T4 Ligase buffer	NEB
T4-PNK	T4 PNK buffer, ATP	NEB
T7-RNA-polymerase	RNAPol Reaction buffer	NEB

2.1.6 Cell lines

Table 12: Cell lines: Cell lines, used in this work, are listed together with their corresponding growth medium and known mutations in the MAPK- and PI3K- pathway (COSMIC cell line project (http://cancer.sanger.ac.uk/cancergenome/projects/cell_lines), Roche Cancer Genome Database (http://rcgdb.bioinf.uni-sb.de/MutomeWeb) and group internal results). For the receipt of the media see table 6.

Cell line	Mutations	Medium
293T	WT	D10
CaCO2	WT	D10
Colo678	KRAS (G12D)	RPMI
HCT116	KRAS (G13D), PIK3CA (H1047R)	D10
HT29	BRAF (V600E), KRAS (Q61L), PIK3CA (P449T)	D10
LIM1215	KRAS (A146T)	D10
RKO	BRAF (V600E), PIK3CA (H1047R)	D10
SW480	KRAS (G12V)	D10

All cells are stored in a cell culture stock in the group of molecular tumor pathology. Cells were checked regularly for their identity using a PCR-based approach and for mycoplasma contamination by the Venor® GeM Mycoplasm Detection kit.

2.1.7 Bacteria

Table 13: Bacteria

Bacteria	**Genotype**	**Company**
DH5α ^{TM}	F– Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK–, mK+) *pho*A *sup*E44 λ– *thi*-1 *gyr*A96 *rel*A1	Invitrogen
PIR1	F– Δ*lac*169 *rpo*S(am) *rob*A1 *cre*C510 *hsd*R514 *end*A *rec*A1 *uidA(ΔMlu I)::pir-116*	Invitrogen
Stbl3^{TM}	F– *mcr*B *mrr hsd*S20 (rB-, mB-) *rec*A13 *sup*E44 *ara*-14 *gal*K2 *lac*Y1 *pro*A2 *rps*L20 (StrR) *xyl*-5 λ- *leu mtl*-1	Invitrogen

2.1.8 Primers and oligonucleotides

The sequences of the primers and oligonucleotides which were used in this work are shown in the supplementary material 8.1. Primers and oligonucleotides were designed using the webtool Primer3 (V 0.4.0).[230] A melting temperature of 60 ℃ was attempted for all primers except the primers used for mutagenesis. In that case a temperature of about 65 ℃ was suggested by the Phusion site-directed Mutagenesis kit.

2.1.9 Plasmids

Table 14: Plasmids: plasmids are shown, which are used for cloning procedures or transfection.

Plasmids	**Function**	**antibiotics conc. in LB**	**Source**
EX-A0114-M02	Vector for transient transfection (AREG)	unknown	GeneCopoeia
pBluescript II KS+	RNA expression vector	Ampicilin 50 ng/µl	kindly provided by M. Morkel, Berlin
pCpGl-basic	CpG-free firefly-Luciferase containing plasmid for promoter analysis	Zeocin 25 ng/µl	kindly provided by M. Rehli, Regensburg
pCpGl-CMV	Positive control for promoter analysis	Zeocin 25 ng/µl	kindly provided by M. Rehli, Regensburg
Renilla	Control plasmid for promoter analysis renilla-Luciferase containing plasmid	unknown	Promega
pCDH-IRES-GFP-Puro	Destination vector	Ampicilin 50 ng/µl	kindly provided by P. Medina Berlin
pDONR^{TM} 221	Donor vector	Kanamycin 50 ng/µl	Invitrogen
pmD2G	Lentivirus packaging vector	Ampicilin 50 ng/µl	kindly provided by A Kramer, Berlin
psPAX	Lentivirus packaging vector	Ampicilin 50 ng/µl	kindly provided by A Kramer, Berlin
EGFP-plasmid	Control vector for transient transfection	unknown	Invitrogen

2.1.10 Software, webtools and databases

Table 15: Software, webtools and databases

Name	Version	Company/Webpage
Human genome	GRCh37/hg19	
Mouse genome	NCBI37/mm9	
Biomart[231]	0.7	http://www.biomart.org/
ClustalW[232]	2.0	http://www.ebi.ac.uk/tools/clustalw2
CTCF-binding sites	unknown	http://bsproteomics.essex.ac.uk:8080
Ensembl Genome Browser	release 68-69	http://www.ensembl.org/index.html
In silico PCR	unknown	http://genome.ucsc.edu/
Ascent Software (Luminoscan)	2.1	Labsystems
Microplate Manager	5.2.1	Bio-Rad
NEBcutter[233]	2.0	http://tools.neb.com/NEBcutter2/
Photoshop	5.0	Adobe
Primer3[230]	0.4.0	http://frodo.wi.mit.edu/
R statistics[234]	2.12.2	http://www.R-project.org/
Sequence Manipulation Suite[235]	2	http://www.bioinformatics.org/sms2/
StepOne^{TM} Software	2.2.2	Applied Biosystems
UCSC Genome browser[236, 237]	unknown	http://genome.ucsc.edu/
WinMDI	2.8	Joe Trotter, The Scripps Institute

2.2 Standard procedures

2.2.1 Standard PCR

Standard PCRs were mainly performed with the AmpliTaq® Gold Taq Polymerase.

Table 16: PCR components and cycling conditions for standard PCR

	Standard PCR
Reaction volume	10 - 50 µl
AmpliTaq® Gold	0.25 - 0.5 U / 20 µl rxn.
Reaction Buffer	1x
dNTPs	25 -200 µM each
Primer forw	0.05 µM - 0.2 µM
Primer rev	0.05 µM - 0.2 µM
Template genomic DNA	10 - 100 ng
Template Plasmids	100 pg

Cycling conditions	95 ℃ 15 minutes
	40x
	{
	95 ℃ 30 seconds
	Tm_{primer} 30 - 60 seconds
	}
	4 ℃ HOLD

Alternatively, standard PCRs were performed using 2x Mastermixes of various suppliers e.g. Multiplex PCR Kit from Qiagen or DreamTaq Green DNA Polymerases from Thermo Scientific. Primer concentrations and template amount as well as cycling conditions remain equal as shown in table 16.

2.2.2 High fidelity PCR

High fidelity PCR was performed either with Phusion® DNA-polymerase or with Pfu Ultra^{TM} High fidelity DNA-polymerase.

Table 17: PCR components and cycling conditions for High fidelity PCR

	PfuUltra^{TM}	**Phusion® HF**
Reaction volume	50 µl	50 µl
DNA-polymerase	2.5 U	1 U
Reaction buffer	1x	1x
dNTPs	1 mM each	0.2 mM each
Primer forw	0.2 µM - 0.5 µM	0.5 µM
Primer rev	0.2 µM - 0.5 µM	0.5 µM
Template genomic DNA	50 -250 ng	50 - 250 ng
Template plasmids	100 pg	100 pg

Cycling conditions	95 ℃ 2 minutes
	40x
	{
	95 ℃ 30 seconds
	Tm_{primer} 30 - 60 seconds
	72 ℃ 1 minute per kb
	}
	72 ℃ 10 minutes
	4 ℃ HOLD
		98 ℃ 30 seconds
		40x
		{
		98 ℃ 10 seconds
		Tm_{primer} +3 ℃ 30 sec
		72 ℃ 30 sec per kb
		}
		72 ℃ 10 minutes
		4 ℃ HOLD

The PCR components and cycling conditions are described in the corresponding section when differing from the description above.

2.2.3 Colony PCR

Besides digestion, a further method to identify a plasmid after transformation is colony PCR. Bacteria colonies were picked and transferred into 2 ml reaction tubes containing 100 µl LB medium. After incubating the tubes at 37 ℃ for two hours standard PCR reactions were performed using primers which identify the plasmid and flank the insert, primers which identify the insert and one plasmid primer plus one insert primer to identify the orientation of the insert. 1 µl of the bacteria suspension was used as template for each reaction.

2.2.4 Strand-specific PCR

Qiagen Multiplex PCR kit was used for strand-specific PCR of bisulfite-converted c-DNAs. Bisulfite conversion was performed using the Epitect-kit (Qiagen) as described by the manufacturer. The PCR-volumes for the strand-specific PCR were set to 10 - 20 µl. Primer concentrations were set to 0.2 µM. The temperature program was as following: 15 minutes at 95 ℃, 40 x (30 seconds at 94 ℃, 60 seconds at 58 ℃, 60 seconds at 65 ℃) and a final elongation step at 72 ℃ for 10 minutes.

2.2.5 Primer phosphorylation

Primers were either purchased phosphorylated at their 5'OH group or they were phosphorylated using T4 Polynucleotide Kinase (T4 PNK). The 10µl reaction contained: 1x T4 PNK buffer (NEB), 2.5 U of T4 PNK, 1 mM ATP and 77.5 µM primer. After 30 minutes at 37 ℃ the enzyme was heat inactivated for 20 minutes at 65 ℃. Phosphorylated primers were stored at -20 ℃ without purification.

2.2.6 Agarose gel electrophoresis

DNA gel electrophoresis To analyze DNA fragment lengths 1.5 - 2 % (w/v) agarose gels were produced as follows: Per 100 ml, 1.5 - 2 g agarose were combined with 100 ml 1x Tris-acetate EDTA buffer. After dissolving by heating in a microwave and cooling to hand warm temperature, 2 µl Ethidium bromide (10 mg/ml) were added and the gel was casted into an appropriate gel chamber. After coagulation 5/6 volume sample plus 1/6 volume 6x loading dye were loaded onto the gel and run for 45 to 90 minutes at 30 V to 70 V in 1x Tris-acetate EDTA buffer. The percentage of the gels as well as the run time and the voltage depended on the expected fragment length. After electrophoresis, gels were applied to the gel documentary system, and fragments were detected by UV excitation of Ethidium bromide, a DNA-double-strand intercalating dye.

RNA gel electrophoresis To analyze RNA-integrity after isolation or to prepare RNA containing membranes for Northern blot experiments, agarose gels were prepared as follows: Per 100 ml, 1.2 g agarose were combined with 72.2 ml RNase-free water. After dissolving by heating in a microwave and cooling to hand warm temperature 10 ml 10x MOPS- buffer and 17,8 ml formaldehyde (37 %) were added. After casting and coagulation of the gel, 1/3 of the sample plus 2/3 of the RNA-loading dye were combined, denatured at 95 ℃ for 2 min and loaded onto the gel. The gel was run at 40 - 50 V for 1 to 4 h. After electrophoresis, gels were applied to the gel documentary system, and fragments were detected by UV excitation of Ethidium bromide, or gels were used for Northern blot experiments.

2.2.7 Plasmid restriction digest

Plasmids and PCR-products to be digested were incubated with the appropriate restriction enzyme and the corresponding buffer (see table 11) and incubated at 37 ℃ for one hour up to over night digestion. Afterwards, the enzymes were heat inactivated when possible and the digest products were isolated when necessary. Digested plasmids were isolated after agarose gel electrophoresis using the kit "QIAquick® Gel extraction kit", whereas PCR products were isolated using the kit "QIAquick® PCR purification kit".

2.2.8 Plasmid *in vitro* methylation

In a 40 µl reaction, 6 µg of plasmids to be methylated were combined with 6 U CpG-Methyltransferase (M.SssI), 160 µM S-Adenosyl-Methionine and 1x NEB buffer 2. *In vitro* methylation was performed at 37 ℃ for 1.5 to 2 h. After heat inactivation of the enzyme at 65 ℃ for 20 minutes, the plasmids were isolated using the kit "QIAquick® PCR purification kit".

2.2.9 Generation of DNA blunt ends

Plasmids or inserts were incubated for half an hour at 37 ℃ with a 10-fold excess of DNA Polymerase I large Klenow fragment to convert sticky ends into blunt ends after digest. Blunt-ended fragments were then either isolated using the "QIAquick® PCR purification kit" or dephosphorylated (2.2.10).

2.2.10 Plasmid dephosphorylation

To reduce plasmid re-ligation, linearized plasmids were incubated for one hour at 37 ℃ with a 10-fold excess of Calf intestinal alkaline phosphatase (CIP). Dephosphorylated plasmids were then isolated using the "QIAquick® PCR purification kit".

2.2.11 Plasmid ligation

T4-ligase was used to ligate linearized dephosphorylated plasmids with inserts prior to transformation. In a 20 µl reaction 50 to 200 ng of a plasmid were combined with an insert at a molar ratio of 1 to 3, 1x ligation buffer and 400 U of T4-ligase. The ratio was calculated using equation 1.

$$\frac{m_{plasmid}\ (in\ ng)\ x\ length_{insert}\ (in\ kb)}{length_{plasmid}\ (in\ kb)}\ x\ ratio(\frac{insert}{plasmid})\ =\ m_{insert}\ (in\ ng) \quad (1)$$

Ligation was mainly performed over night at 16 ℃. As controls, a ligase free ligation and an insert free ligation was added to the experiment.

2.3 Cell culture

2.3.1 Thawing of cells

Cells are stored in liquid nitrogen. First the cryovial was removed from the storage unit and set into a 37 ℃ water bath. After fast thawing, the cells were transferred into 15 ml reaction tubes. The appropriate cell culture medium was added dropwise to the cell suspension to limit the probability of an osmotic shock. Afterwards, the reaction tube was centrifuged at 800 rpm for 5 min. Then the supernatant was replaced by fresh medium, the pellet was resuspended and transferred to a clean cell culture flask and stored ON in a cell incubator at 37 ℃ and 5 % CO_2. On the next day, the medium was again replaced by fresh medium to remove dead detached cells from the culture flask.

2.3.2 Maintenance

The cell lines grew in 75 cm^2 cell culture flasks at 37 ℃ supplemented with 5 % CO_2 and approximately 96 % H_2O. Cells were splitted when the cell monolayer covered approximately 70 to 90 % which occurred every second or third day.

The standard splitting procedure was as follows: After the medium was fully removed from the cells, the monolayer was rinsed one time with 1x PBS (see table 6). Afterwards, 1.5 ml Trypsin/EDTA-solution (see table 5) were added to the monolayer to detach the cells from the flask. Detaching reaction was performed in the cell incubator at 37 ℃ for 5 to 10 minutes. The completely detached cells were dissolved in 8.5 ml medium and transferred into a new 15 ml reaction tube followed by 5 min centrifugation at 800 rpm. After complete removal of the supernatant the cell pellet was redissolved in 10 ml fresh medium. 1 to 5 ml of the cell suspension, depending on the time, when a full flask was needed the next time, was then refilled into the cell culture flask and supplemented with fresh medium to get 10 ml total volume. Alternatively, the cells were frozen or seeded appropriately for an experiment.

2.3.3 Freezing

To store cells for future experiments, cells were frozen in liquid nitrogen. After splitting, the cell pellet was resuspended in 1.5 ml medium supplemented with 10 % DMSO and transferred into a cryovial. To avoid destruction of the cells by the formation of ice crystals a slow decrease in temperature was necessary. The vial was put into a cryobox filled with IPA and stored first in a -80 ℃ freezer. After the vial was completely frozen, it was transferred into the liquid nitrogen hold.

2.3.4 Cell treatment procedures

During this work several colorectal cancer cell lines were treated with several different compounds. These experiments were done either in 96-well plates with a cell growth

area of 0.31 cm^2 per well, in 24-well plates with 1.9 cm^2, 6-well plates with 9 cm^2 or 10 cm petri dishes with 60 cm^2. The 96-well plates were used for growth determination experiments (XTT), the 24-well plates were used for protein phosphorylation analysis and the other two devices enabled the collection of supernatant for protein analysis or RNA for RNA-expression analysis.

In general, four timepoints were analyzed per experiment. Three wells of a 96-well plate were used per treatment for growth determination, one well of a 6-well plate or one petri-dish was used per treatment for RNA and supernatant sample collection per timepoint. The timepoints analyzed for cells treated with EGF-receptor inhibitors were 24, 48, 72 and 96 h. For cells treated with epigenetic interfering compounds, the timepoints tested included 72, 96, 120 and 144 h, because these compounds can only interfere during cell cycle and therefore are supposed to be effective after a longer period.

The number of cells seeded per well or plate depended on the overall duration of the experiment and the cell growth rate to avoid overgrowth. The number of seeded cells was approximately 2500 to 5000 cells per cm^2 growth area for the shorter experiments up to 96 h, whereas the number of seeded cells was approximately 1000 to 4200 cells per cm^2 growth area for the longer experiments up to 144 h. After seeding, the cells incubated at 37 ℃ for 24 h. Then treatment was performed daily by replacing the complete medium with fresh medium or fresh medium containing the compound or the solvent. The concentrations of the compounds used and the appropriate solvents are shown in table 18.

Table 18: Compounds in treatment experiments

Name	Type	Solvent	Concentration
Cetuximab	EGFR inhibitor	NaCl 9 g/l	10 µg/ml
Erlotinib	EGFR inhibitor	DMSO	10 µM
Gefitinib	EGFR inhibitor	DMSO	10 µM
DAC	DNMT inhibitor	NaCl 9 g/l	2.5 µM
Cambinol	HDACi	DMSO	20 µM
SAHA	HDACi	MetOH	1 µM
TSA	HDACi	DMSO	25 ng/ml
Valproat	HDACi	MetOH	1 mM, 2mM
Zebularine	DNMT inhibitor	NaCl 9 g/l	100 µM

2.3.5 Cell post-treatment experiments

After treating the cells for 144 h with DNA methyltransferase inhibitors (DNMT inhibitors) or HDACis the cells were replated into 96-well plates as follows. The supernatant was removed and stored for ELISA. Afterwards, the plates or wells were rinsed with 1x PBS (see table 6) and treated with 1 ml of Trypsin/EDTA solution (see table 5). After 5 minutes of incubation at 37 ℃ the cells were redissolved in 5 ml fresh medium and centrifuged for 5 minutes at 800 rpm. Then the pellet was redissolved in

fresh medium and cells were counted using a Neubauer improved cell counting chamber or the TC10 cell counter. 1500 cells per well of untreated, solvent treated and compound treated cells were replated into four 96-well plate, reflecting four different timepoints in a cell proliferation assay (Cell Proliferation Kit II (XTT)).

After 24 h of incubation the supernatant of each well was replaced by medium or medium plus EGFR inhibitors (see table 18) or the appropriate solvent. For the next four days cellular viability was measured as described by the manufacturer. XTT absorbance was measured 4 h and 24 h after dispension of the XTT solution. Also for the next four days the medium of the remaining plates was replaced daily by fresh medium or medium plus inhibitor/solvent.

2.3.6 Cell transfection experiments

To transfect cells with DNA using Lipofectamine^{TM} 2000, cells were seeded into a suitable well that they were 70% - 90% confluent at the time of transfection. In table 19 the composition of each transfection reaction is shown for the appropriate well size, as used during this work.

Table 19: Compounds in transfection experiments

	6 well	**12 well**	**96 well**
Solution 1			
Optimem	250 µl	100 µl	8.3 µl
DNA	1 µg - 2µg	250 ng	21 ng - 55 ng
Solution 2			
Optimem	250 µl	100 µl	8.3 µl
Lipofectamine^{TM} 2000	9 µl	4 µl	0.33 µl

Solution 1 and 2 were incubated at RT for 5 minutes before they were mixed. After a second incubation step for 20 minutes at RT, the transfection solution was applied to the cells dropwise. The cells were then incubated at 37 ℃ in the cell incubator for four hours. Afterwards, the medium was completely replaced by fresh medium. The cells were then put into the cell incubator until progression of the experiment.

2.3.7 Sample collection

To detect proteins in the supernatant of cells, the medium was transferred into a new reaction vessel. Detached cells were pelleted by centrifugation at 10000 rpm and 4 ℃. The supernatant was then transferred into a new reaction tube and stored for ELISA-experiments at -20 ℃ (see table 7). To analyze the mRNA, the cells were processed as described in section 2.6.

2.4 Standard bacteria procedures

2.4.1 Generation of chemically competent bacteria PIR1

One commercial chemically competent bacteria stock PIR1 was used to generate up to 200 chemically competent bacteria stocks for transformation of pCpGl-basic and pCpGl-basic derived plasmids. Bacteria from the commercial stock were transferred into 6 ml LB-medium and agitated at 37 ℃ over night. On the next morning the complete volume was transferred into 1.4 l LB-medium and agitated for four hours at 30 ℃. Then the bacterial suspension was centrifuged for 25 minutes at 4000 rpm and 4 ℃. The pellet was then dissolved in 150 ml TFBI -buffer (see table 6) and incubated on ice for one hour. As a next step the suspension was again centrifuged for 25 minutes at 4000 rpm and 4 ℃. Finally, the pellet was dissolved in 20 ml TFBII -buffer (see table 6) and aliquoted in 1.5 ml reaction tubes which were chilled with dry ice. The 100 µl aliquots were then used directly for transformation or stored at -80 ℃.

2.4.2 Agar plates and transformation

Agar plates: LB-Agar was melted slowly in a microwave. 15 ml per agar plate were then transferred into 50 ml reaction tubes, cooled down to hand warm temperature and combined with antibiotics (see table 14). The LB-Agar was then casted into 10 cm petri-dishes and cooled down to room temperature. For PIR1 transformed bacteria low-salt agar plates were casted. For bacteria, transformed with pBluescript II KS+, the agar plates were supplemented with X-Gal (80 µg/ml) and IPTG (200 µM) to enable blue-white screening.

Transformation: 25 µl of commercial chemically competent bacteria stocks (see table 13) or 100 µl of self made bacteria stocks (see 2.4.1) were combined with 2 to 10 µl ligated plasmids and incubated on ice for 30 minutes. Then the bacteria were incubated at 42 ℃ for 30 seconds (heatshock) and combined with 200 µl SOC-medium. After a one hour agitation step at 37 ℃, 100 µl were streaked out on an agar plate and incubated at 37 ℃ over night. Besides the plasmids of interest, every time an undigested isolated plasmid was transformed as a positive control, an empty bacteria stock was used as a negative control, a ligase free ligation was used as a molecular cloning procedure control.

2.4.3 Plasmid identification and isolation

After transformation, plasmid containing bacteria could be identified through their resistance towards the antibiotics on the agar plate. Colonies formed on the plate. Using a pipet tip, a number of these colonies was picked and transferred into 2 ml LB-medium, supplemented with the appropriate antibiotic. For pCpGl-basic derived plasmids, colonies were picked into 2 ml low-salt LB medium. Only insert-containing pBlue-

script II KS+ -plasmid containing white colonies were picked and transferred into 2 ml LB-medium. Each 2 ml sample was then agitated at 37 ℃ over night.

On the next day 1.5 ml of each sample was used for plasmid isolation using the "Qiaprep Spin Miniprep Kit". Isolated plasmids were then tested for the insert and the orientation of the insert by restriction digest (see 2.2.7), or they were characterized by colony PCR (see section 2.2.3) using insert and plasmid specific primers. The remaining 0.5 ml were shock frozen as described in section 2.4.4. Alternatively, after identification of positive clones, the remaining bacteria were transferred into 100 ml or 500 ml LB-medium or low-salt LB medium supplemented with the corresponding antibiotic (see table 14) and agitated over night at 37 ℃. On the next day an aliquot of the bacteria was shock frozen as described in section 2.4.4. The remaining bacteria were used for plasmid isolation. Plasmids were isolated using either the kit "Qiagen® Plasmid Midi kit" or the kit "Qiagen® Plasmid Maxi kit". Isolated plasmids were then checked again by restriction digest and by Sanger sequencing.

2.4.4 Storage of bacteria

0.5 to 1 ml of bacteria suspension was transferred into 2 ml cryovials and combined with an equal amount of glycerin stock solution. After mixing, the vial was shock frozen using liquid nitrogen. The bacteria were stored at -80 ℃. If necessary, bacteria could be re-amplified by transferring a small amount of the frozen stock using a pipet tip into fresh LB-medium supplemented with the corresponding antibiotic.

2.5 Protein expression analysis

Proteins were detected by ELISA experiments using the kits described in table 7.

2.5.1 Protein isolation

Proteins were isolated using 300 µl lysis buffer per 10 cm petri dish. Lysis buffer contained 78 % Homogenization buffer, 20 % cOmplete, Mini, +EDTA (1 tablet per 2 ml ddH_2O) and 2 % Triton X-100. After removal of cell medium and two wash steps with PBS, lysis buffer was added, cells were scraped using a rubber policeman and transferred into a 2 ml reaction tube. The tube was sonicated for one minute and stored on ice for ten minutes. This step was repeated two times. Afterwards, the tube was centrifuged for ten minutes at 13000 rpm and 4 ℃. The supernatants were stored at -20 ℃. Protein concentrations were determined prior to ELISA using the BCA protein assay kit.

2.5.2 Enzyme-linked Immunosorbent Assay (ELISA)

The protocol was performed mainly as described by the manufacturer. In brief, a 96-well Nunc Maxisorp® plate was coated with capture antibodies, covered with an adhesive strip

and stored over night at RT. On the next day, each well was washed 3 times with ELISA Wash solution and blocked with 300 µl Reagent diluent for at least 1 hour. After each well was washed again 3 times, 100 µl sample or standard was dispensed in duplicates into the wells. As sample, either 100 µl cell supernatant or 10 µg cell lysate was applied to each well. Cell lysates and standards had been diluted prior dispense using ELISA Reagent diluent. The plate was covered with an adhesive strip and incubated over night at 4 ℃. On the next day, the plate was washed again 3 times and detection antibodies were dispensed into the wells. After incubation at RT for at least 2 hours, the wells were washed again 3 times, incubated for 20 minutes with Streptavidin-HRP, washed 3 times and incubated with 100 µl TMB per well. After 15 to 20 minutes a blue color occurred indicated the abundance of the protein of interest. A color change to yellow occurred after dispensing 100 µl 1N H_2SO_4 per well. Optical density of each well was measured at 450 nm with a correction wavelength of 540 nm. Standard curves were generated in Excel using a four parameter logistic curve fit. Please refer to the manual of the appropriate ELISA-kit for exact concentrations of the antibodies and the standards used, as well as the dispensed volumes (see table 7). The results were evaluated by comparing the optical densities of each well with the optical density of the standard curve. For cell lysate measurement, the protein amounts in pg refer to an input amount of 10 µg total protein. For supernatant measurement, the results were normalized to the RNA concentrations of the corresponding samples (see section 2.6.1). One pg RNA equals one pg produced per day per 1 ng isolated RNA.

2.6 Gene expression analysis

2.6.1 RNA isolation

RNA was isolated either using the kit "High Pure RNA isolation kit" (Roche Diagnostics) or using the kit "RNeasy® Mini Kit" (Qiagen). After the supernatant was removed from the cells, the plate or the well was rinsed with 1x PBS and the cells were collected with the corresponding kits' lysis buffer. When using the Qiagen kit, the lysed cells were then homogenized using the kit "Qiashredder" (Qiagen). The remaining procedure of RNA isolation was performed as described in the manual of the kit. Isolated RNA was measured using a Nanophotometer (IMPLEN) and stored at -80 ℃.

2.6.2 CDNA synthesis

The kit "Transcriptor High fidelity cDNA synthesis kit" (Roche Diagnostics) was used for cDNA synthesis. Either 1 µg or the maximum possible volume of isolated RNA was converted. CDNA synthesis was performed using random hexamers and cDNA was stored afterwards at -20 ℃. For real-time-PCR, cDNA was used without any further purification.

2.6.3 Real-time PCR using the TaqMan approach

Real-time-PCR was mainly performed using the TaqMan® approach from ABI. Here, probes labeled with a fluorescent dye were used to observe the PCR amplification. Each 10 µl reaction was composed of 1x TaqMan® Gene Expression Mastermix, one or two 1x primer probe sets and the cDNA template. The 1x TaqMan® Gene Expression Mastermix contained the Taq-polymerase, dNTPs and cofactors in unknown concentrations. Each 1x primer probe set contained the locus-specific primers and the TaqMan® probes, labeled either with the dye FAM^{TM} or VIC®. The template was applied directly from the cDNA-synthesis or diluted up to 1:10. For gene expression analysis the signal of the gene of interest was compared with the signal of an unregulated control gene. In this work *UBE2D2* was used as a control. The temperature program for real-time PCR was as follows:

10 minutes - 95 ℃
40x
{ 10 seconds - 95 ℃, 1 minute - 60 ℃, data collection }

Due to their capability of analyzing two probe sets per reaction, duplex reactions were possible, when the appropriate singleplex reactions gave similar results. After verification experiments (data not shown) the pairs *AREG/UBE2D2* and *EREG/UBE2D2* were confirmed to be used in duplex reactions. In singleplex reactions, the ΔCt values (Ct (control) - Ct (gene of interest)) were calculated from the means of the Ct values of three technical replicates. In multiplex reactions, each ΔCt value could be calculated per well. Statistical analysis, calculation of ΔΔCt and relative quantification was performed as described by ABI.[238]

2.6.4 Real-time PCR using the SybrGreen approach

SybrGreen® is a further means of analyzing mRNA-expression. In contrast to TaqMan® an intercalating dye is used to observe the amplification of the PCR-product. Each 20 µl reaction was composed of 1x Power SybrGreen® Mastermix (Applied Biosystems) or 1x GoTaq® qPCR Mastermix (Promega), 0.2 µM of forward and reverse primers and the cDNA-template. The 1x commercial mastermixes contained the Taq-polymerase, dNTPs and cofactors in unknown concentrations. The template was applied directly from the cDNA-synthesis or diluted up to 1:10. As for TaqMan, *UBE2D2* was used as control gene. The temperature program for real-time PCR was as follows:

10 - 15 minutes - 95 ℃
40x
{ 15 seconds - 95 ℃, 1 minute - 60 ℃, data collection }
As a final step, a melting curve was performed (0.5 ℃ per minute) to analyze for unspe-

cific PCR-products.

The ΔCt values (Ct (control) - Ct (gene of interest)) were calculated from the means of the Ct values of three technical replicates. Statistical analysis, calculation of ΔΔCt and relative quantification was performed as described by ABI.[238]

2.7 Protein phosphorylation analysis

Protein phosphorylation was analyzed using the Bio-Plex® Suspension Array System (Bio-Rad). This method is a protein multiplex detection method, involving two specific antibodies per target protein in a sandwich ELISA-based approach. One antibody is coupled to a polystyrene-bead filled with two fluorescent dyes in distinct target specific ratios. Up to 100 different ratios are distinguishable, meaning that theoretically 100 different proteins might be detected. The other antibody is labeled with biotin which is necessary to couple a streptavidin-coupled fluorescent dye (phycoerythrin) to the antibody. The beads bound to the target proteins are then subjected to a flow cell, where two lasers excite the beads individually. One laser excites the polystyrene-bead, which identifies the target, the other laser excites the dye at the second antibody. The measured fluorescence can be used to quantify the target molecules per bead.

2.7.1 Protein isolation

Cells were grown and treated in 24-well plates. The medium was removed and the cells were washed 1x with cold 1x PBS. 25 µl Bio-Plex® cell lysis buffer supplemented with Bio-Plex® factor 1 (1:250), factor 2 (1:500) and PMSF (500 µM) were dispensed into each well. After sealing, the cells were shook on ice on a shaker for 20 minutes and afterwards the plates were frozen at -80 ℃. After thawing, the lysis buffer was transferred into reaction tubes and protein concentration was determined by the BCA protein assay kit.

2.7.2 Data aquisition

1. The first bead-coupled antibodies of either the pEGFR assay (Singleplex) or the pMEK1, pERK1/2 and pAKT assays together (Multiplex, see also table 8) were diluted 1:100 in Bio-Plex® wash buffer. 50 µl of these dilutions were transferred per well into a 96-well filterplate.
2. The liquid was removed using a plate vacuum manifold and the plate was washed three times with 80 µl Bio-Plex® wash buffer per well.
3. 10 to 15 µg of isolated protein were transferred into the wells, the plate was sealed, followed by an incubation step at RT over night on a plate shaker (400 rpm).
4. The second step was repeated once.
5. The second detection antibodies of either the pEGFR assay (Singleplex) or the pMEK1, pERK1/2 and pAKT assays together (Multiplex) were diluted 1:50 in Bio-

Plex® detection buffer. 25 µl of these dilutions were transferred per well into the plate.
6. After a further incubation step on the plate shaker for 45 minutes (400 rpm), the plate was washed again as indicated in step 2.
7. 50 µl 1x Streptavidin-PE was transferred into each well and the plate was incubated in the dark for 10 minutes on the plate shaker (400 rpm).
8. The plate was washed again as indicated in step 2, but using Bio-Plex® resuspension buffer instead of wash buffer.
9. After the last wash step, 85 µl Bio-Plex® resuspension buffer was inserted in each well and the plate was subjected to the Bio-Plex® 200 machine.
10. 100 beads per target protein were set in the mode "Run at high RP1-target". The upper gate value was set to 14,000.

2.8 Functional analysis of genomic regions

2.8.1 Plasmid generation for functional analysis

Figure 1:

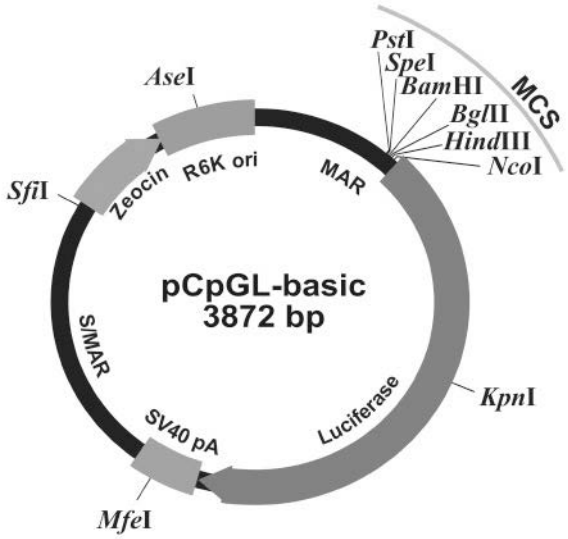


Figure 1: pCpGl-basic: Klug and Rheli[239] designed a CpG-free vector for promoter analysis experiments. This vector contained a multiple cloning site in front of a firefly Luciferase gene. Various constructs were cloned into the multiple cloning site. The figure was adopted from: Klug and Rehli. Functional analysis of promoter CpG methylation using a CpGfree luciferase reporter vector. Epigenetics. 2006; 1(3): 127-130

PCpGl-basic, a CpG-free Luciferase containing vector (Klug and Rheli,[239] see figure 1), was used to analyze CpG-methylation dependent functional aspects of the *AREG* promoter and the exon 2.

Human genomic DNA was used as input for a standard PCR containing primers which harbor locusspecific parts and 5'- overhang sequences. In the overhang sequence

the forward primers contained a *BamHI* restriction site whereas the reverse primers contained a *HindIII* restriction site. The primer sequences are shown in the supplementary material 8.1. They were used to amplify the AREG promoter sequence, the AREG exon 2 and the exon 2 sequence divided into two parts. The temperature program of the PCR was as follows:

15 minutes - 95 ℃
10x
{30 seconds - 95 ℃, 30 seconds - 60 ℃, 30 seconds - 72 ℃}
30x
{30 seconds - 95 ℃, 30 seconds - 70 ℃, 30 seconds - 72 ℃}
HOLD 4 ℃

After PCR the PCR products as well as pCpGl-basic were digested with *BamHI* and *HindIII* as described in 2.2.7. Isolated linearized pCpGl-basic was then dephosphorylated (see 2.2.10), ligated with the isolated PCR product (2.2.11), transformed into PIR1, identified and isolated (2.4.2 and 2.4.3). Due to the *BamHI* and *HindIII* restriction sites, plasmids were generated with forward oriented inserts.

For the generation of the promoter-containing plasmids, pCpGl-basic was digested with *HindIII* and *NcoI*. The CpG-free EF1-promoter was amplified by PCR from the control plasmid pCpGl-CMV-EF1 using the same program as shown above. Primers were used having 5' overhangs which contain either a *HindIII* restriction site (forward primer) or an *NcoI* restriction site (reverse primer). To create a CpG-free Tata-binding motif, two oligonucleotides were used, which anneal to a DNA double strand Tata motif containing sequence with 3'-overhangs on both strands. The 3' overhangs could be ligated into digested *HindIII* and *NcoI* restriction sites. The EF1 -PCR product as well as the Tata-binding motif forming oligonucleotides were ligated with the linearized plasmid to form pCpGl-Tata and pCpGl-EF1. Plasmids were transformed and isolated (see 2.4.2 and 2.4.3).

PCpGl-basic, pCpGl-Tata, pCpGl-EF1 and pCpGl-AREG-promoter were used to create plasmids harboring the exon 2 insert sequence in both orientations. First, the plasmids were linearized with *BamHI*. The exon 2 insert as well as the exon 2 parts were created as described above and digested with *HindIII* and *BamHI*. For the exon 2 reverse plasmid (pCpGl-AREG-exon2-R), the insert was created directly using primers without 5' overhang. The single-stranded overhangs of both the linearized plasmids as well as the digested PCR products were filled to doublestranded DNA using the Klenow fragment (see 2.2.9). The plasmids were then dephosphorylated and ligated with the PCR products (see 2.2.10 and 2.2.11). Plasmids were transformed and isolated as described in section 2.4.2 and 2.4.3.

Finally, plasmids, harboring specific mutations were necessary to analyze the influence of the CTCF-binding site on the function of the exon 2 region. pCpGl-AREG-exon2-

AREG-promoter, pCpGl-AREG-exon2-R-AREG-promoter and pCpGl-AREG-exon2-R-Tata were mutated using the kit "Phusion site-directed mutagenesis kit" and the 5' phosphorylated primers CTCFtoHINDIII_1/2 (see: 2.2.5). These primers bind downstream the CTCF-binding site on both strands of the plasmid and have a 5'- overhang that form a *HindIII* restriction site (see figure 2). The formation of PCR products was performed as described by the manufacturer. The PCR-temperature program was as follows.

30 seconds - 98 ℃
40x
{10 seconds - 98 ℃, 30 seconds - 71 ℃, 2 minutes - 72 ℃}
10 minutes - 72 ℃
HOLD 4 ℃

Circular plasmids were created by ligation. Afterwards, the plasmids were transformed into PIR1 bacteria, isolated and characterized as described in 2.4.2 and 2.4.3.

All generated plasmids are listed in supplementary section 8.2. The sequences of all plasmids are shown in supplementary section 8.3.

For methylation dependent promoter function analysis, several plasmids needed to be *in vitro* methylated. Methylation was performed as described in section 2.2.8. Methylated plasmids were analyzed for their methylation by restriction digest using the methylation sensitive restriction enzyme *AvaI*. One example, on how methylated plasmids were identified, is shown in supplementary section 8.5.

Figure 2:

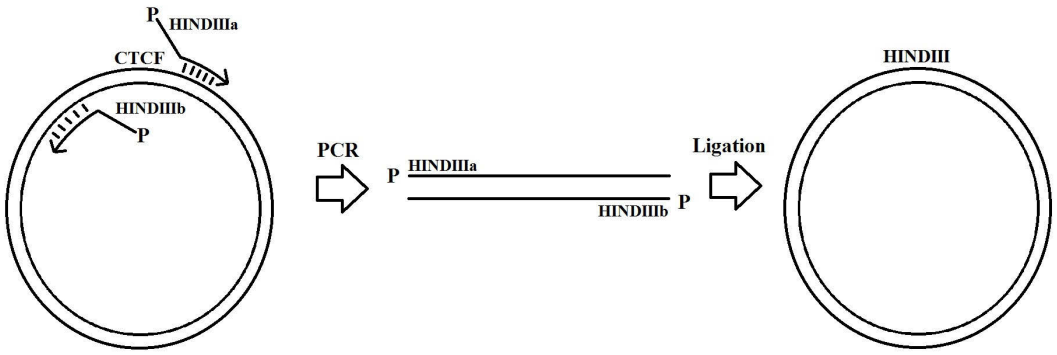


Figure 2: Principle of plasmid mutagenesis: Phosphorylated primers which bind downstream a CTCF-binding site on both strands of a plasmid and which harbor sequence unspecific parts, forming together a *HindIII* restriction site (HindIIIa, HindIIIb) were used in the "Phusion site-directed mutagenesis kit" to create linear products to form new plasmids after ligation.

2.8.2 Cell transfection and sample preparation

Plasmids were transfected either in two replicates in 12 well plates or in three replicates in a 96 well Costar Cluster-Plate as shown in section 2.3.6. All pCpGl-basic derived plasmids, which contain a *Firefly Luciferase* gene were transfected simultaneously with a *Renilla Luciferase* gene containing control plasmid. After replacing the transfection medium with fresh medium, cells were incubated in the cell incubator over night. On the next day the medium was completely removed from the plate and each well was washed once with 1x PBS.

12 well plate assay 200 µl 1x Passive Lysis Buffer (PLB: Dual-Luciferase® Reporter (DLR^{TM} Assay) were pipetted into each well. Cells were scraped from the plate and transferred into a new reaction tube. The tubes were stored at -80 ℃.

96 well plate assay 20 µl 1x PLB (Dual-Luciferase® Reporter Assay) were pipetted into each well. The plate was shaken for 30 minutes at 600 rpm on the plate shaker IKA MTS 2/4, sealed and afterwards stored at -80 ℃.

2.8.3 Detection of Firefly and Renilla Luciferase activity

Luminescence measurement using Lumat LB 9507 20 µl of PLB -lysed cells were transferred into a 5 ml polypropylene tube (Sarstedt) and mixed with 90 µl of Luciferase Assay Reagent II (LARII: Dual-Luciferase® Reporter (DLR^{TM}) Assay). Firefly luminescence was measured for 10 seconds. Afterwards, 90 µl of 1x Stop & Glow reagent was tranferred into the glas vial and mixed carefully. Renilla luminescence was then measured for 10 seconds.

Luminescence measurement on Luminoscan RS The Costar Cluster-Plate, used for transfection, was thawed to room temperature and shaken for 10 minutes at 600 rpm. Alternatively, 20 µl of PLB -lysed cells were transferred into the wells of a 96-well Costar Cluster-Plate. Firefly and Renilla luminescence were measured as follows:

1. Start a clock. Pipet 70 µl of Luciferase Assay Reagent II into the first well of the plate
2. Wait 10 seconds.
3. Pipet 70 µl of Luciferase Assay Reagent II into the second well of the plate
4. Wait 10 seconds.
5. Repeat that procedure until a maximum of 8 wells were pipetted.
6. Load the plate into the plate reader.
7. When 2 minutes are passed from the first pipet step, start measuring luminescence.
8. Measure each well for 10 seconds in high gain and collect the data.
9. When 4 minutes are passed from the first pipet step, start measuring again.

10. Measure each well for 10 seconds in high gain and collect the data.
11. When 6 minutes are passed from the first pipet step, start measuring again.
12. Measure each well for 10 seconds in high gain and collect the data.
13. Remove plate from the reader, stop the clock and repeat the whole procedure for the same wells but pipetting Stop & Glow reagent instead of LARII.
14. Repeat all steps, to measure all wells on the plate.

The described procedure ensured that every well was measured at the same time after pipetting the substrate solution and that each well was measured 3 times independently.

2.9 AREG protein over-expression experiments

2.9.1 Transient transfection

LIM1215 and HCT116 cells were transfected with plasmids encoding AREG (EX-A0114-M02, GeneCopoeia) and with a plasmid encoding EGFP as a control. Transfection was performed in 6 well plates as described in section 2.3.6. AREG protein expression was tested by ELISA 24, 48, 72 and 96 hours after transfection.

2.9.2 Lentiviral transduction

Lentivirus generation Several steps are required to generate lentiviruses for the over-expression of AREG:
1. 5' ATTB-site containing primers were used to amplify the AREG transcript sequence by high fidelity PCR (2.2.2). The plasmid EX-A0114-M02 (see table 14) was used as a template. In addition to the complete AREG transcript, a shorter transcript lacking the C-terminal end of the AREG protein was created. The C-terminus has been previously described as a potential regulatory protein in the nucleus.[28]
2. The kit: "Gateway® Technology with Clonase^{TM} II" was used to produce the Entry clones by BP-reaction. As a donor vector 150 ng of pDONR^{TM} were recombined with 50 fmol of the PCR product. The recombined plasmid was transformed into DH5α ^{TM} as described in section 2.4.2. Isolated plasmids were sequenced with M13 primers prior to the next step.
3. The expression plasmids were produced by LR recombination reaction using 150 ng of the entry clones and 150 ng pCDH-IRES-GFP-Puro as described in the kit "Gateway® Technology with Clonase^{TM} II". Recombined plasmids were transformed into Stbl3, selected and isolated as described in section 2.4.2 and 2.4.3. The correct inserts were identified by digest with *NheI* and *HindIII* and comparison with *in silico* data (see supplementary material section 8.6).
4. Viruses were produced using 293T cells at S2 conditions. Transfection of the viral components and the expression clone was performed using the CalPhos^{TM}-kit from Clontech in 293T cells at about 40% confluency. 8.4 µg of the expression clones or as

control pCDH-IRES-GFP-Puro were combined in 526 µl H_2O with 6 µg psPAX and 3.6 µg pmD2G plasmids. The overall volume was adjusted to 600 µl with 2M Calcium solution (supplied with the kit). The plasmid mixture was added dropwise to 600 µl 2x HBSS (supplied with the kit) while vortexing carefully. After an incubation of 20 minutes at RT, the mixture was spread dropwise onto the 293T cells. The cells were incubated over night in the cell incubator. On the next day the medium was replaced by fresh medium. The following day the virus containing supernatants were collected and centrifuged at 800 rpm for 5 minutes to remove cellular debris. The procedure was repeated again on the next day. All supernatant was then passed through a 0.2 µl filter and stored at -80 ℃.

Transduction LIM1215 cells were used for transduction at 70 % confluency. First the medium was replaced by medium containing protamine sulfate (c = 8 µg/ml) followed by a 30 minute incubation step at 37 ℃. Afterwards, the cells were transduced by adding the filtered virus containing supernatants to the wells. Medium was changed each day for the next two days. On the second day the fresh medium was supplemented with puromycin (c = 0.02 µg/ml) to select for the transduced cells. Selection was performed for two weeks. The cells were then stored as described in section 2.3.3 or used for experiments.

2.10 Northern blot analysis

To detect a potential *AREG* reverse non-coding RNA, Northern blot analyses were performed with RNA-probes from different regions within the *AREG* gene.

2.10.1 Probe generation

RNA probes were derived from the plasmid pBluescript II KS+ , which harbors a T3- and T7-RNA promoter flanking a multiple cloning site. Inserts were cloned into the plasmid in the following way:
1. The inserts were produced by high fidelity PCR (2.2.2) containig the Pfu Ultra Polymerase and phosphorylated primers (2.2.5). The Tm of the reaction was set to 58 ℃.
2. pBluescript II KS+ was linearized with *EcoRV* to create blunt ends (2.2.7). Afterwards, the blunt ends were dephosphorylated (2.2.10).
3. After agarose gelelectrophoresis (2.2.6) the linear plasmid was isolated (QIAquick® Gel extraction kit, Qiagen).
4. Insert and plasmid were ligated (2.2.11) and transformed into DH5α ^{TM} (2.4.2).
5. pBluescript II KS+ enables blue-white screening for colonies containing plasmid plus insert. Only white colonies were isolated and characterized (2.4.3). The sequences are shown in supplementary material 8.8.
6. Isolated plasmids were digested either with *EcoRI* or *HindIII* to obtain linear plas-

mids for RNA-polymerization (see supplementary material 8.7).
7. 1 µg of the digested plasmids were introduced to the kit "T7 High Yield RNA Synthesis kit". Probes were generated as described in the manual using either T7 or T3 RNA-polymerase, and [α-^{32}P]-UTP. Incubation was performed at RT for 2 h.
8. The template DNA was digested using DNaseI by incubating the reaction for 15 minutes at 37 ℃.
9. Finally the probes were purified using the Illustra MicroSpin G50-columns (GE-Healthcare) and radioactivity was determined using a radioactivity counter.

Steps 7 to 9 were performed by the technician Cornelia Gieseler.

2.10.2 Blot-membrane generation

The technician Cornelia Gieseler generated the blot-membranes. Equal amounts of poly-A-RNA, which were isolated using the kit "PolyATtract® mRNA Isolation System IV" (Promega) were loaded per lane onto an RNA-agarose gel (see 2.2.6) Gels were run at 50 V for 4 h. The RNA-transfer from the gels to the blots was performed in subsequent steps:
1. A basin was filled with 20x SSC -buffer and covered with a glas-plate.
2. A Whatman 1MM paper was wet with 20x SSC and placed on the glas plate in a way that two sites of the paper reach into the 20x SSC in the basin.
3. Two gel-sized Whatman 1MM papers were placed onto the wet Whatman paper.
4. After wetting in 5x SCC for 5 min, the gel was placed upside-down onto the Whatman papers.
5. A nylon transfer membranes (Hybond N+) was placed onto the gel.
6. Then, 4 dry 1 MM Whatman papers were placed on top of the nylon membrane.
7. Paper towels were put on top of the Whatman paper to a height of approx. 10 cm and a final glas plate was placed on top of the towels.
8. The blotting device was adjusted with a heavy weight and left over night.
9. On the next day, all towels and papers were removed, the samples were marked on the membrane and the membrane was put in 2x SCC.
10. The membrane was dried for 1 minute by placing it on a dry Whatman paper.
11. The membrane was transferred with the RNA side-down on a UV transilluminator (254 nm wavelength), and the RNA on the membrane was crosslinked for an appropriate length of time.
12. Finally, the membrane was shrink-wrapped and stored at -20 ℃.

2.10.3 Hybridization and readout

Hybridization and readout was also performed by the technician Cornelia Gieseler using the following steps:
1. The membranes were unwrapped and transferred into hybridizing tubes.

2. 10 ml Ultrahyb -Buffer (Amerhsham) was added to each membrane for pre-hybridization at 68 ℃ for 2 hours
3. Then, each RNA-probe was added completely to the appropriate membrane and the membranes were incubated at 68 ℃ ON.
4. On the next day, the hybridization buffer plus probes were removed and the membranes were washed using buffer N1 for 15 minutes at 30 ℃.
5. Afterwards, the membranes were washed again using buffer N2 for 15 minutes at 68 ℃.
6. Then the membranes were shrink-wrapped and put into a film-chamber together with 2 radiographic films.
7. After ON-incubation at -80 ℃, the films were developed by using the developer and fixer system from Kodak.

2.11 Immunohistochemistry

The IHC-experiments were performed by the technician Kerstin Möhr using 0.3 µm paraffin slices on standard microscopic glass slides.

2.11.1 Dewaxing

1. The slides were incubated at 70 ℃ for 20 minutes.
2. The slides were incubated 3 times for 5 minutes in 100 % Xylene.
3. The slides were incubated 2 times for 5 minutes in 100 % Ethanol.
4. The slides were incubated each 3 minutes in 96 %, 90 %, 80 % and 70 % Ethanol.
5. Finally, the slides were incubated 3 minutes in H_2O.

2.11.2 Demasking

1. The slides were put into citrate buffer and incubated in a steam cooker for 20 minutes.
2. The slides were cooled down slowly to room temperature and incubated for 15 minutes in H_2O.
3. Finally, the slides were incubated for 10 minutes in a 3 % H_2O_2-solution.

2.11.3 Blocking and 1st antibody incubation

1. 100 to 150 µl Peroxidase-Blocking Solution, Dako Real^{TM} was pipetted onto the cells on the slides.
2. The slides were incubated for 30 minutes in a wet chamber.
3. After rinsing with 1x TBS, the AREG-Antibody was diluted 1:400 in antibody Diluent (Zytomed) and transferred to the cells.
4. The slides were incubated over night at 4 ℃ in a wet chamber.

2.11.4 2nd antibody incubation and staining

1. After rinsing with 1x TBS, the 2nd antibody (rabbit anti goat, HRP-conjugated) was pipetted onto the slides in a 1:350 dilution.
2. The slides were incubated 1 hour at room temperature.
3. After rinsing with 1x TBS, Liquid DAB + Substrate Chromogen System (Dako) was added to the slides.
4. The antibody-staining was stopped after 60 seconds by rinsing with H_2O.
5. The slides were incubated for 10 seconds in hematoxylin-solution to stain the nuclei.
6. The slides were washed in H_2O.
7. Finally, the slides were incubated for each 10 seconds in 70 %, 80 %, 90 %, 96 % and 100 % Ethanol.

2.11.5 Embedding

1. 20 to 75 µl of 1:1 solution containing Xylene and Vitro-Clud was pipetted onto the slides.
2. A cover-glass was put onto the slide and the slide was dried.

2.12 External methods

Methylation analysis as well as Xenograft experiments were performed by external co-laborators within the Colonet consortium. Methylation analysis was performed by Dr. Sascha Tierling in the Institute of Genetics and Epigenetics in Saarbrücken, Xenograft experiments were performed by Maria Rivera in the group of experimental pharmacology in the Max-Delbrück Center Berlin-Buch.

2.12.1 Methylation analysis

The following protocols were written and operated by Dr. Sascha Tierling:

DNA preparation and bisulfite treatment DNA was prepared using standard protocols. Bisulfite treatment was performed on 500 ng genomic DNA. Briefly, DNA was treated with 2 M sodium bisulfite and 0.6 M NaOH, then denatured for 15 min at 99 ℃ and incubated for 30 min at 50 ℃. We introduced two thermo spikes of 99 ℃ for 5 min followed by two incubation steps of 1.5 h at 50 ℃. Purification was achieved by loading, desulfonation with 0.3 M NaOH and washing with 1 x TE on a microcon YM-30 column (Millipore, Schwalbach, Germany). Bisulfite DNA was eluted in 50 µl 1 x TE.

Promoter methylation analysis by 454 GS-FLX sequencing Amplicons were generated using region- specific primers with the recommended A and B adaptors (Roche, Mannheim, Germany) at their 5' -end. PCR of bisulfite-treated DNA was performed in

a total volume of 30 µl containing 3 µl Hot Star-Taq reaction buffer (Qiagen, Hilde, Germany), 2.4 µl dNTPs, 2.5 mM each, 100 nmol of each primer, 1.25 U Hot Star-Taq (Qiagen, Hilden, Germany) and 2 µl of bisulfite DNA. After initial denaturation at 95 ℃ for 15 min, 40 cycles were carried out (denaturation 95 ℃ for 1 min, annealing 54 ℃ for 1 min, extension 72 ℃ for 1 min and final extension 72 ℃ for 5 min). PCR products were visualized on 1.2 % agarose gels, purified using the Gel/PCR DNA Fragments extraction kit (AVEGENE, Taipei, Taiwan) and measured by intercalating fluorescence dye (Qubit HS-Kit, Invitrogen, Darmstadt, Germany) using Qubit Fluorometer (Invitrogen, Darmstadt, Germany). After amplicon pooling, emulsion PCR was performed using Lib-A emPCR protocols. DNA-containing beads were recovered, enriched and sequenced from the A-adaptor on a XLR70 TitaniumPicoTiterPlate according to the manufacturers protocols (Roche, Mannheim, Germany).

MsSNuPE-experiments and HPLC separation Fifty ng of genomic DNA were used as a template in a 30 µl reaction volume in the presence of 3 mM Tris-HCl, pH 8.8, 0.7 mM $(NH_4)_2SO_4$, 50 mM KCl, 2.5 mM $MgCl_2$, 0.06 mM of each dNTP, 3 U HotFire DNA polymerase (Solis BioDyne, Tartu, Estonia) and 1 µM primers. 5 µl of the PCR products were treated with 1 µl of ExoSAP (1:10 mixture of Exonuclease I and Shrimp Alkaline Phosphatase, USB) for 30 min at 37 ℃. To inactivate the ExoSAP enzymes the reaction was incubated for 15 min at 80 ℃. Afterwards, 14 µl primer extension mastermix (50 mM Tris-HCl, pH 9.5, 2.5 mM $MgCl_2$, 0.05 mM of all four ddNTPs, 3.6 µM of each SNuPE primer, 2.5 U Termipol DNA polymerase (Solis BioDyne, Tartu, Estonia)) was added. Primer extension reactions were performed at 96 ℃ for 2 min followed by 50 cycles 96 ℃ / 30 sec, 50 ℃ / 30 sec, 60 ℃ / 2 min. Separation of products was conducted at 50 ℃ by continuously mixing buffer B (0.1 M TEAA, 25 % acetonitril) to buffer A (0.1 M TEAA), either over 15 min: 23-31 % (AREG A2), or 15 min: 22-35 % (EREG A2).

2.12.2 Xenograft experiments

The *in vivo* experiments (generation and treatment of xenografts, as well as tissue collection) were carried out by Maria Rivera in the group of experimental pharmacology in the Max-Delbrück Center Berlin-Buch. $5x10^6$ LIM1215 cells were injected subcutaneously into immunodeficient NMRI nu/nu mice. Per treatment five mice were used. Three independent experiments were performed. In the first experiment (MV10107) the influence of Valproat alone and in combination with Erlotinib on tumor growth was examined. In the second experiment (MV10532) the treatment mode was changed to evaluate the effect of Valproat treatment prior to Erlotinib treatment on tumor growth. And in the third experiment, the influence on tumor growth of 5-Azacytidine alone and in combination with Erlotinib was tested. After cell injection, tumor growth was measured two times a week. When the tumors reached a volume of approximately 0.1 to 0.2 cm^3 the

treatment started, which was for MV10107 on day 19, for MV10532 on day 12 and for MV10533 on day 13 post injection. The treatment modes and the concentrations of the used inhibitors are shown in table 20. The concentrations correspond to the highly tolerable doses in human treatment. After finishing the experiments, the mice were killed and the tumors were shock frozen in liquid nitrogen. Afterwards, the tumors were stored at -80 ℃. Parts of the tumors were used for RNA isolation and AREG/EREG mRNA expression analysis (see section 2.2.7) as well as immunohistochemistry experiments.

Table 20: Design of xenograft experiments to compare *in vitro* results in an *in vivo* system: i.p.: intraperitoneal injection, p.o.: oral treatment

MV10107	**Application**	**Days**	**Dose (mg/kg/inj.)**
Solvent (PBS)			
Valproat	i.p.	Mon-Fri, 2x/day	200
Erlotinib	p.o.	Mon -Fri	50
Cetuximab	i.p.	every 7th day 2x (q7dx2)	50
Valproat + Erlotinib	i.p. p.o.	Mon-Fri, 2x/day Mon -Fri	200 50

MV10532	**Application**	**Days**	**Dose (mg/kg/inj.)**
Solvent (NaCl)			
Solvent (NaCl) Erlotinib	i.p. p.o.	day 1 to 6 day 7 to 11	 50
Erlotinib + Valproat	p.o. i.p.	day 1 to 6 day 7 to 11	50 200
Valproat + Erlotinib	i.p. p.o.	day 1 to 6 day 7 to 11	200 50

MV10533	**Application**	**Days**	**Dose (mg/kg/inj.)**
Solvent (NaCl)			
Erlotinib	p.o.	days 1-5, 7, 13-17, 22, 23	50
5-Azacytidine	i.p.	days 1-5, 7, 13-17, 22, 23	5
5-Azacytidine + Erlotinib	i.p. p.o.	days 1-5, 7, 13-17, 22, 23 days 1-5, 7, 13-17, 22, 23	5 50

3 Results

3.1 *Amphiregulin* and *Epiregulin* are differentially expressed in colorectal cancer cell lines

Figure 3:

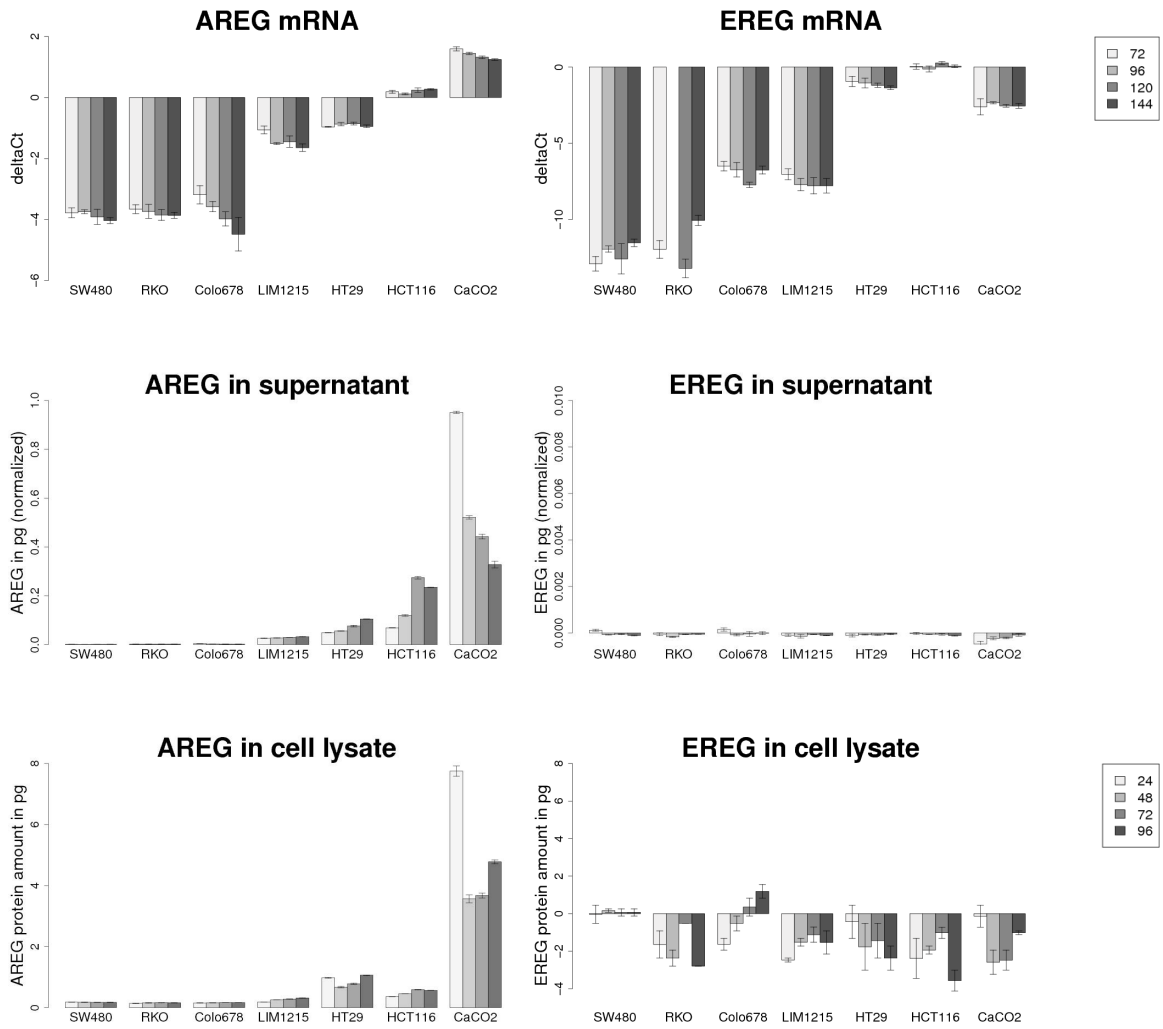


Figure 3: *AREG* and *EREG* expression in different colorectal cancer cells: Seven colorectal cancer cell lines were tested for their *AREG* and *EREG* mRNA expression and protein levels at 4 different timepoints. 1st row: The ΔCt values were calculated by subtraction of the *AREG* or *EREG* Ct-value from the control Ct-value. 2nd row: The AREG or EREG protein amounts in the supernatant were normalized to the corresponding RNA concentration. 3rd row: The AREG and EREG protein amounts are shown per 10 µg total cell lysates (see section 2.5.2).

The mRNA expression of *Amphiregulin* and *Epiregulin* was tested in several colorectal cancer cells at four timepoints by real-time PCR. The amounts of the corresponding proteins in the supernatant or the cell lysates were also tested by ELISA experiments (see figure 3). EREG protein was neither detected in the cell lysates nor in the supernatants by the ELISA assay used. Negative values in the ELISA-barplots

indicate that the ELISA raw data of these samples was equal or lower than the raw data of the negative controls. Therefore, although comparison with the standard curve led to negative values, these values should be considered as zero. In contrast to EREG, there was a good correlation between *AREG* mRNA expression and protein amounts. Low *AREG* mRNA expression correlated with a lack of detectable AREG protein in the supernatant or in the cell lysate, whereas high *AREG* mRNA expression correlated with high AREG protein amounts. Although there was a discrepancy between the AREG protein amount in HCT116 cells between the supernatant and the cell lysate, the data could still be used to divide the cells in three groups of *AREG* mRNA and protein expressing cells and two groups of *EREG* mRNA expressing cells. The AREG groups consisted of cells with low *AREG* mRNA expression and no or only very low detectable amounts of AREG protein (SW480, RKO and Colo678), cells with medium *AREG* mRNA expression and low detectable amounts of AREG protein (LIM1215 and HT29) and cells with high *AREG* mRNA expression and high amounts of AREG protein (HCT116 and CaCO2). The EREG groups consisted of cells with no or very low *EREG* mRNA expression (SW480, RKO, Colo678 and LIM1215) and cells with medium to high *EREG* mRNA expression (HCT116, CaCO2).
The *EREG* mRNA expression correlated with the *AREG* mRNA expression, too. The cells comprising the low AREG expressing group appeared in the group of no or very low *EREG* mRNA expressing cells (SW480, RKO, Colo678). In a similar way, the cells comprising the high AREG expressing group appeared in the group of medium to high *EREG* mRNA expressing cells (HCT116, CaCO2). However, this only holds true for the *EREG* mRNA expression, since EREG proteins were not detected.

3.2 *AREG* and *EREG* are regulated epigenetically

When testing HCT116 cells for their mRNA expression patterns with and without expressed DNA methyltransferases, Sers *et al.* found out that the *AREG* mRNA expression was reduced in the cells, lacking DNMTs.[240] These data indicated that at least AREG is regulated by epigenetic regulatory mechanisms. To address this, two cell lines of the low AREG expressing group (SW480, RKO), one cell line of the medium AREG expressing group (LIM1215) and two cell lines of the high AREG expressing group (HCT116, CaCO2, see section 3.1) were tested for their AREG and EREG expression after treatment with the DNA methyltransferase inhibitor 5-aza-2'-deoxycytidine (Decitabine, DAC) and Zebularine or after treatment with histone deacetylase Inhibitors (HDACi).
DAC is a derivative of Deoxycytidine whereas Zebularine is a derivative of Cytidine. Both are precursors of Deoxycytidine triphosphate, a building block for DNA. After incorporation into DNA, DAC as well as Zebularine inhibit DNA methyltransferases and as a consequence lead to a demethylation of the whole genome after several rounds of

DNA replication.

Histone deacetylases are responsible for the formation of heterochromatine by removing acetyl groups from lysines at the N-termini of histones. As a consequence the affinity of histones to DNA increases, which leads to a blocking of the DNA for transcription factors and to a down-regulation of the transcriptional machinery. By inhibiting histone deacetylases, transcriptional rates are increased all over the genome.

3.2.1 *AREG* and *EREG* expression increases after treatment with the DNMT inhibitor DAC

Figure 4:

120 h: 144 h:

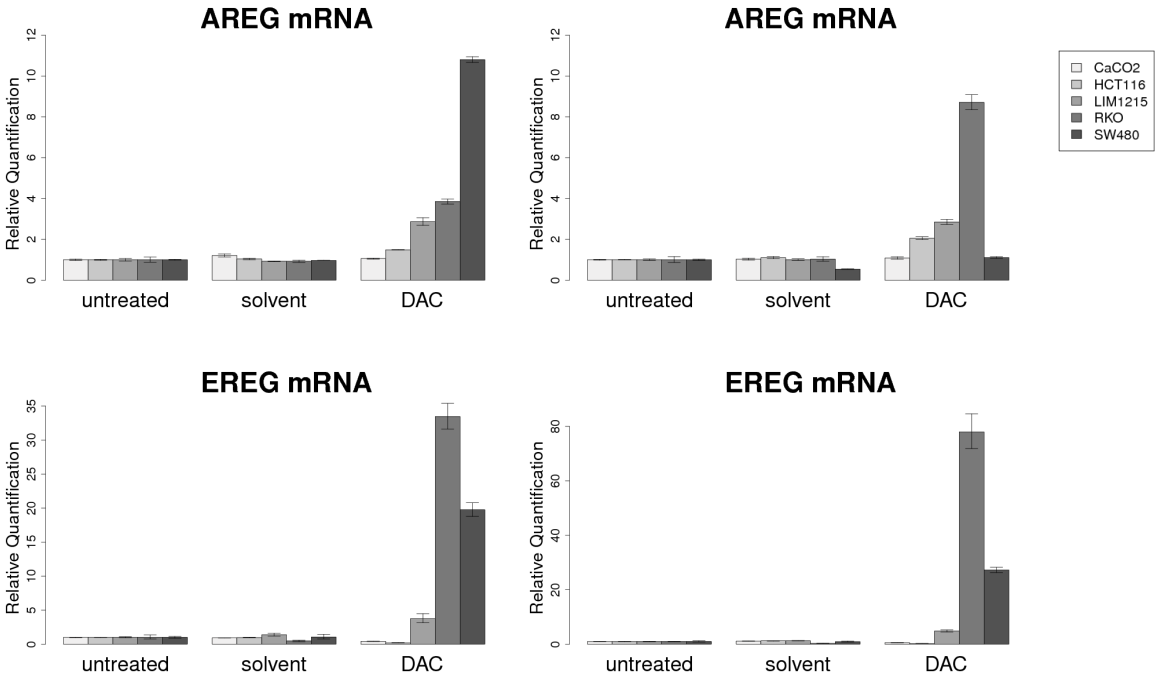


Figure 4: *AREG* and *EREG* mRNA expression after DAC treatment: Five cell lines were tested for their *AREG* and *EREG* mRNA expression after treatment with 2.5 µM DAC for 120 and 144 h. The data were sorted according to the *AREG* mRNA expression of the untreated cells (see figure 3). In each plot the relative quantification of the expression is shown compared to the untreated control (RQ = 1).

Figure 5:

120 h: 144 h:

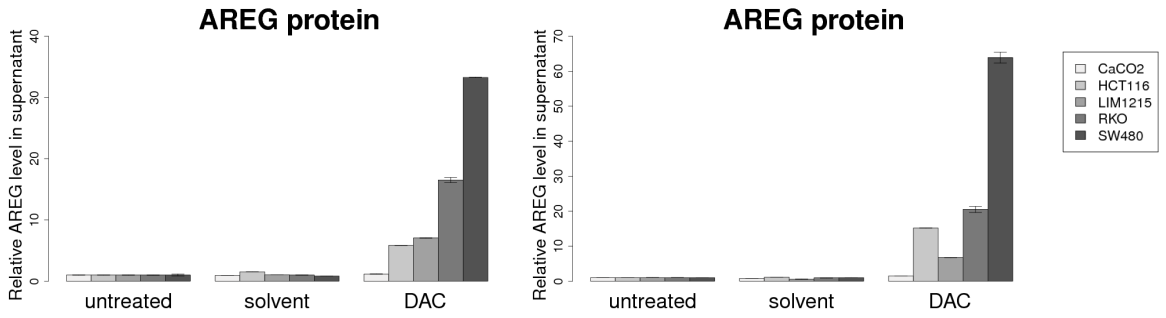


Figure 5: AREG protein expression after DAC treatment: Five cell lines were tested for their AREG protein amount in the supernatant after treatment with 2.5 µM DAC for 120 and 144 h. The data were sorted according to the AREG mRNA expression of the untreated cells (see figure 3). In both plots the relative AREG level is shown compared to the untreated control (rel. AREG = 1).

As shown in figures 4 and 5, DAC led to an up-regulation of *AREG* mRNA and AREG protein as well as *EREG* mRNA. EREG protein was not tested. By comparing the initial AREG and EREG levels (see figure 3) and the effect of DAC onto mRNA and protein expression (see figures 4 and 5), it was observed that the increase of *AREG* and *EREG* mRNA and protein levels by DAC was higher in cells which expressed lower initial levels of AREG and EREG and vice versa. One exception was SW480 at 144 h showing a reduced *AREG* mRNA expression but an increased protein level.

The cell lines were also treated with 100 µM of the DNA methyltransferase inhibitor Zebularine. But compared to DAC, Zebularine did not change the AREG expression after 120 or 144 h of treatment (see supplementary material 8.9, figures 73 and 74). Neither AREG mRNA expression nor AREG protein levels in the supernatant were increased.

3.2.2 *AREG* and *EREG* expression increases after treatment with histone deacetylase inhibitors

Figure 6:

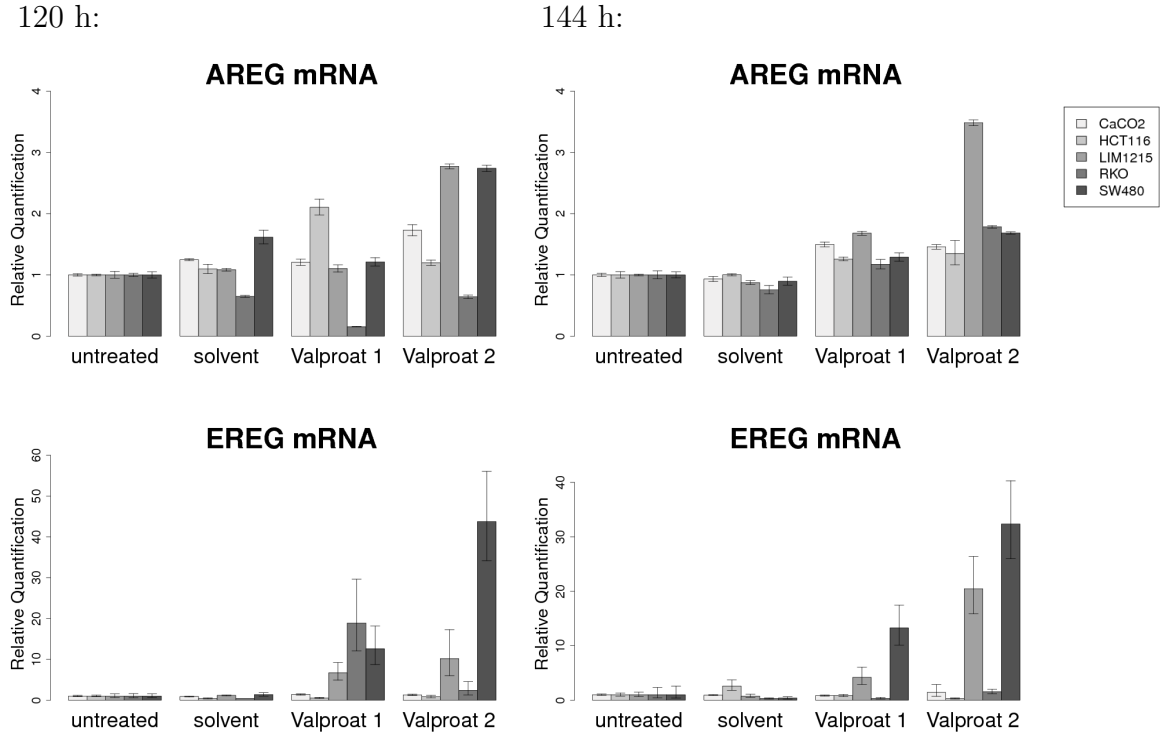


Figure 6: *AREG* and *EREG* mRNA expression after Valproat treatment: Five cell lines were treated with 1 mM or 2 mM Valproat (Valproat 1 or 2) for 120 and 144 h. The data were sorted according to the *AREG* mRNA expression of the untreated cells (see figure 3). In each plot the relative quantification of the expression is shown compared to the untreated control.

Figure 7:

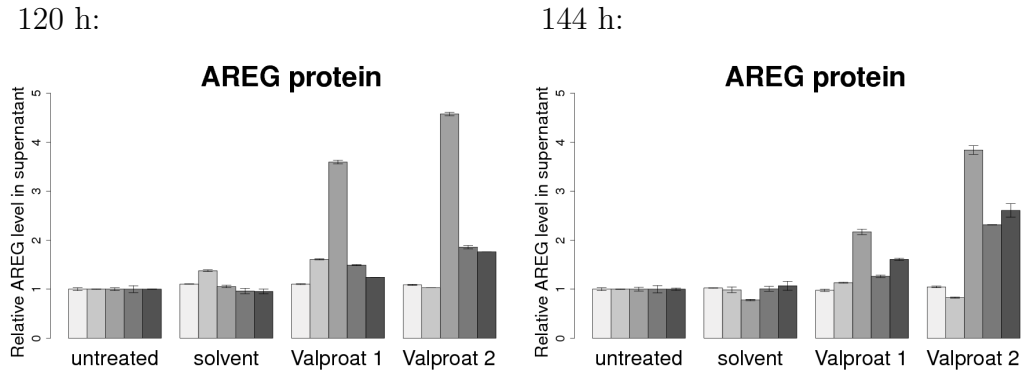


Figure 7: AREG protein expression after Valproat treatment: Five cell lines were treated with 1 mM or 2 mM Valproat (Valproat 1 or 2) for 120 and 144 h. The data were sorted according to the *AREG* mRNA expression of the untreated cells (see figure 3). In both plots the relative AREG level is shown compared to the untreated control.

To test if *AREG* and *EREG* expression can also be changed by HDACis, the cells

were treated for 120 and 144 h with 1 and 2 mM Valproat as described in section 2.3.4. As shown in figure 6 and 7, there is no correlation between the *AREG* and *EREG* expression in untreated cells and the change of *AREG* and *EREG* expression by Valproat, as seen for DAC treated cells (see figure 4). Treatment with 2 mM Valproat increased the *AREG* mRNA expression only in LIM1215 by a factor of three to four at both timepoints. For the other cells, Valproat seemed to affect the expression in a different way. For SW480 cells, an increased *AREG* mRNA expression was also seen. Compared to the untreated control, an increase by a factor of 2.5 was seen after 120 h treatment. However, at 144 h *AREG* mRNA expression declined again to 1.5 fold. Valproat affected the *EREG* mRNA expression to a larger extent. For the very low expressing cells LIM1215 and SW480 treatment with 1 and 2 mM Valproat increased the *EREG* mRNA level between 5 and 60 fold at both timepoints. In contrast, RKO only showed an increased *EREG* mRNA level after 120 h, when treated with 1 mM Valproat. The EREG protein expression after Valproat treatment was tested by Stephan Bartels (Charité Berlin, Institute for Pathology) in his Master thesis.[241] But similar to the results obtained for untreated cell lines in this work (see figure 3) EREG protein was not detected in any of the tested samples, irrespective of Valproat treatment.

Valproat treatment influenced the AREG protein levels (figure 7) similar to the *AREG* mRNA expression. LIM1215 increased levels up to 5 times. However, unlike *AREG* mRNA expression, 2 mM of Valproat led to an increase of AREG protein in RKO and SW480 after 144 h of treatment. Thus, mRNA expression and protein levels are not consistent following inhibition of histone deacetylation.

Figure 8:

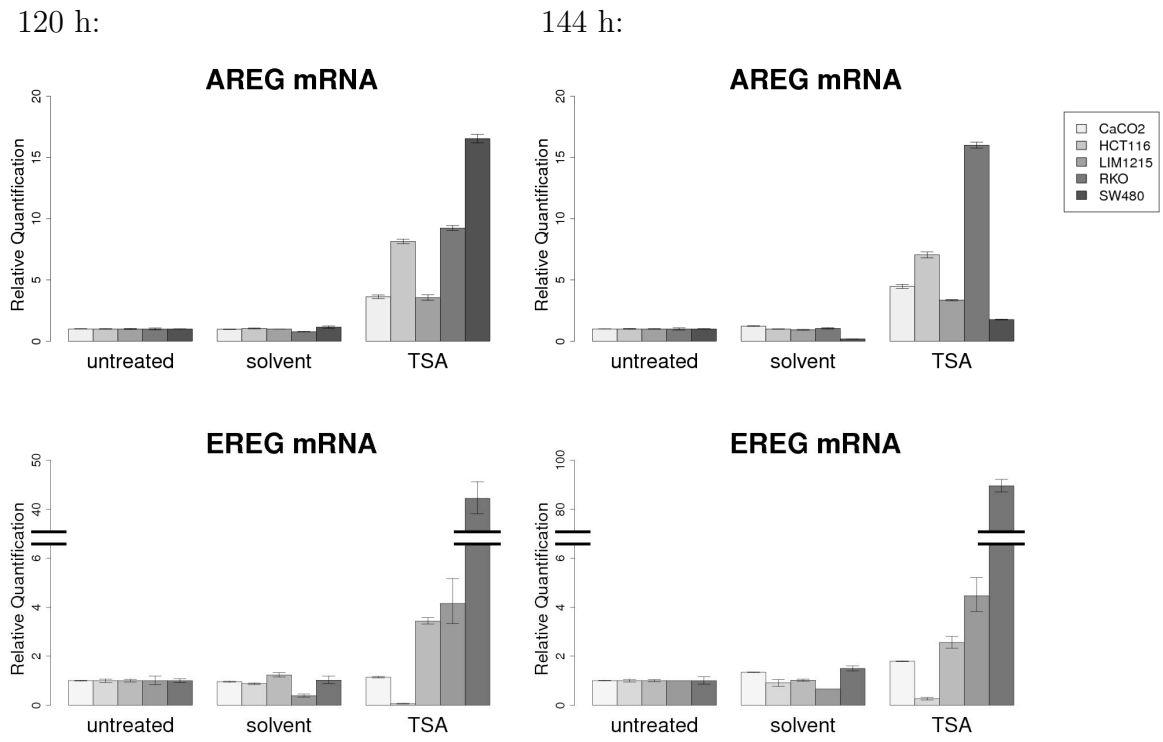


Figure 8: *AREG* and *EREG* mRNA expression after TSA treatment: Five cell lines were treated with 25 ng/ml TSA for 120 and 144 h. The data were sorted according to the *AREG* mRNA expression of the untreated cells (see figure 3). In each plot the relative quantification of the expression is shown compared to the untreated control.

Figure 9:

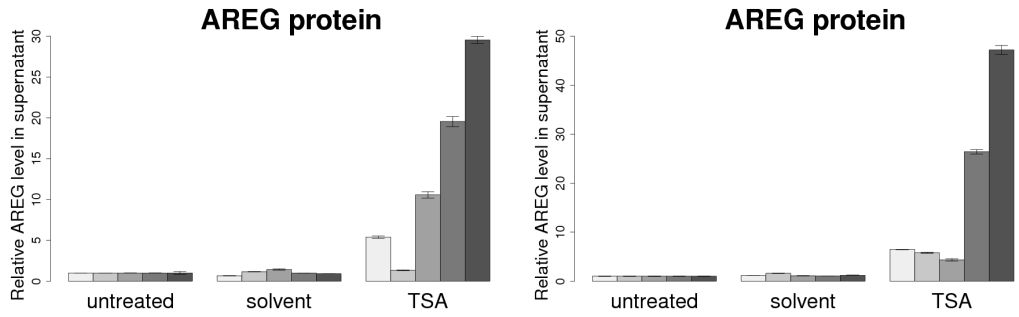


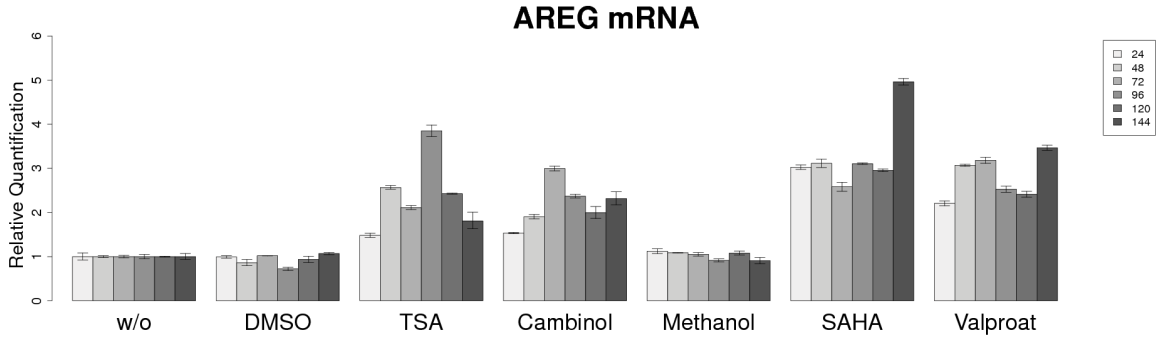
Figure 9: AREG protein expression after TSA treatment: Five cell lines were treated with 25 ng/ml TSA for 120 and 144 h. The data were sorted according to the *AREG* mRNA expression of the untreated cells (see figure 3). In both plots the relative AREG level is shown compared to the untreated control.

To confirm the influence of histone deacetylation on *AREG* or *EREG* expression, the effect of other HDACis was tested. Trichostatin A (TSA) led to an up-regulation of the *AREG* mRNA level in all tested cells after treatment for 120 h and in all cells except SW480 cells after 144 h (see figure 8). However, in that cell line *EREG* mRNA

expression increased by about 90 fold compared to the untreated control. For the other cell lines except HCT116 cells the *EREG* mRNA level was also increased upon treatment with TSA. AREG protein levels increased also for all cells but in contrast to the mRNA result, the strongest effect was seen in SW480 cells (see figure 9).

Figure 10:

LIM1215:



SW480:

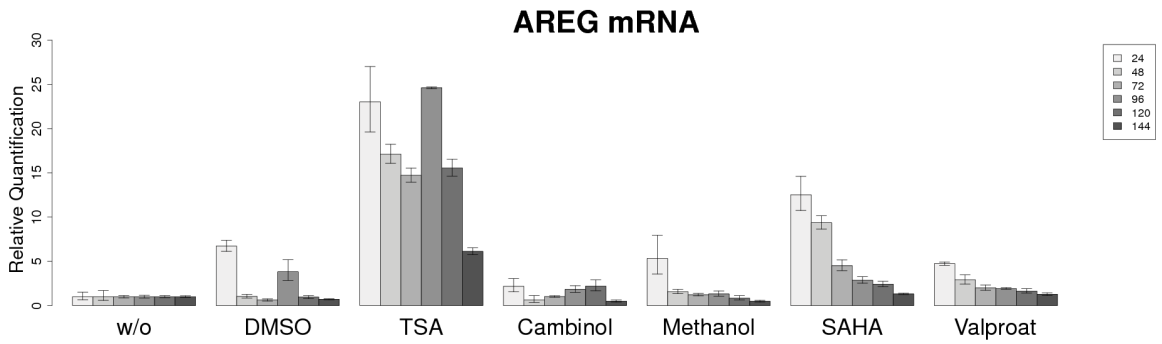


Figure 10: *AREG* mRNA expression after treatment with different HDACis: LIM1215 and SW480 were treated with different HDACis for 144 h (TSA: 25 ng/ml, Cambinol 20 µM, SAHA 1 µM, Valproat 1 mM, DMSO: solvent control for TSA and Cambinol, Methanol: solvent control for SAHA and Valproat). In both plots the relative quantification of the expression is shown compared to the untreated control.

To investigate, if inhibition of histone deacetylases generally leads to an up-regulation of *AREG*, and to address, if the abrupt changes in the AREG expression are common in the cells, LIM1215 and SW480 cells were tested for their *AREG* mRNA expression after treatment with the previously tested TSA and Valproat, but also with two additional HDACis, SAHA and Cambinol. *AREG* mRNA was analyzed for 6 time-points between 24 h and 144 h. In figure 10 is shown that all four HDACis in LIM1215 and all HDACis except Cambinol in SW480 led to an up-regulation of *AREG* mRNA. In contrast to LIM1215, the HDACis led to a higher up-regulation of *AREG* mRNA expression in SW480 at early timepoints, but the effect diminished towards the late timepoints, although the cells were treated daily with the HDACis. In LIM1215 the *AREG* mRNA expression remained almost stable at all timepoints.

To conclude, *AREG* and *EREG* expression were mainly increased by the DNA methyltransferase inhibitor DAC in different cells. The effect depended on the initial *AREG* or *EREG* expression. Expression was also increased by HDACi treatment but in this case it did not depend on initial *AREG* or *EREG* expression. In contrast to expression changes by DAC treatment, expression changes by treatment with HDACis strongly varied between the timepoints of the experiments and between different experiments, too. Also discrepancies between *AREG* and *EREG* mRNA and protein expression were seen after treatment. Nevertheless, the experiments show that AREG and EREG might be regulated by epigenetic mechanisms. In the next step, the potential epigenetic regulatory control regions of both genes were identified.

3.3 *AREG* and *EREG* promoters are mainly unmethylated in colorectal cancer cell lines

The *AREG* gene as well as the *EREG* gene are located at chromosome 4q13.3. The promoter regions of both genes were tested first for their methylation by second generation 454 GS-FLX sequencing. The locations of the promoters in relation to the genes are shown in figure 11.

Figure 11:

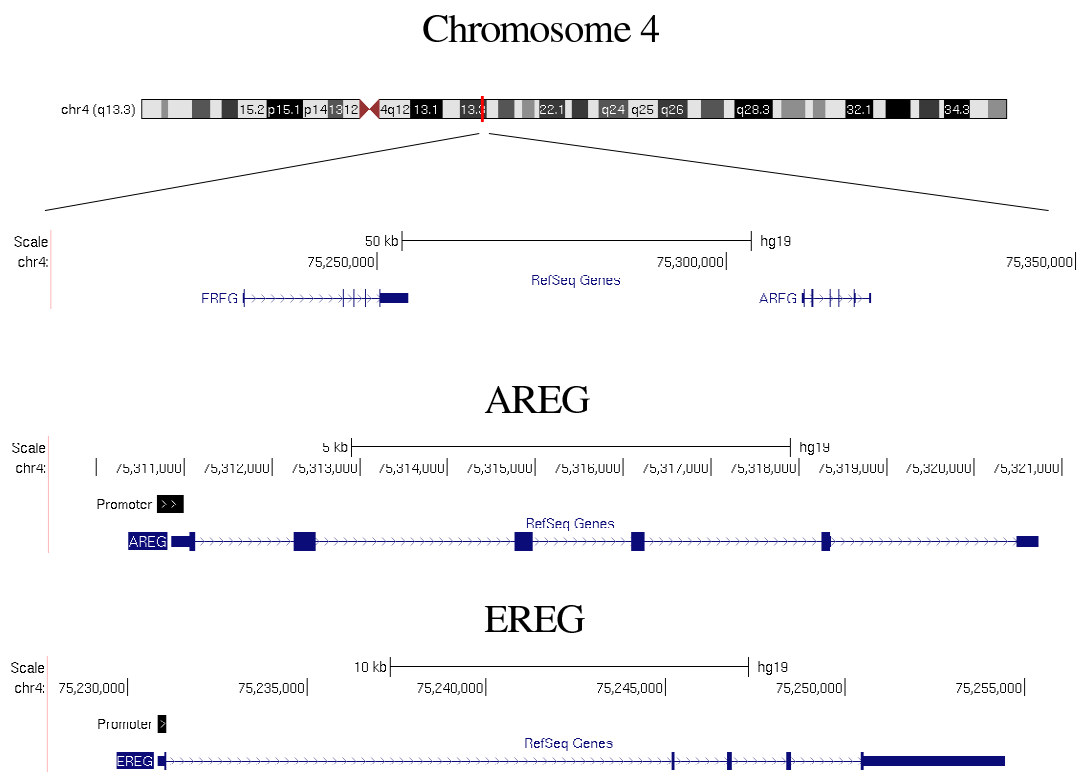


Figure 11: Chromosomal positions of *AREG* and *EREG* genes and gene promoters: Both genes *AREG* and *EREG* are located on chromosomal band 4q13.3. The positions of the corresponding promoters are indicated by black bars.

The experiments leading to figure 12 and the figure itself were made by Dr. Sascha Tierling (Institute of Genetics and Epigenetics, University of Saarland, Saarbrücken).

Figure 12:

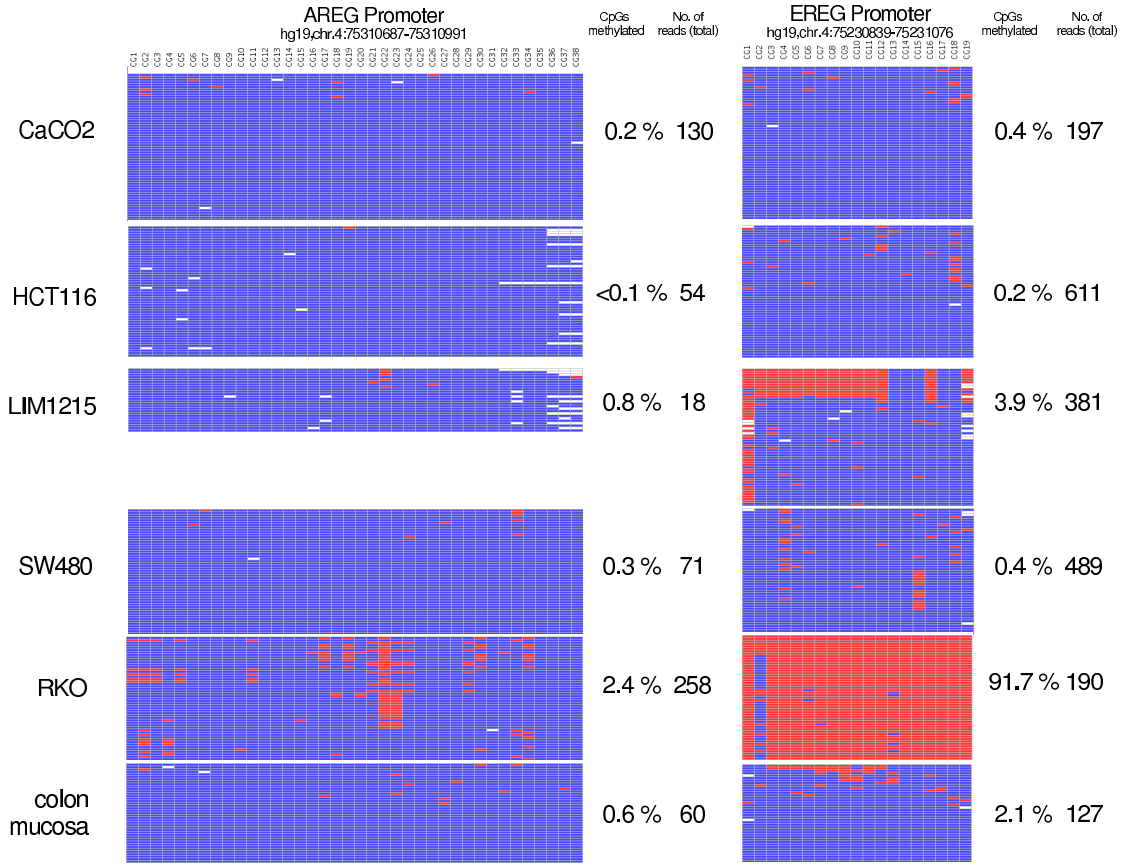


Figure 12: Methylation of *AREG* and *EREG* promoters after 454 GS-FLX deep bisulfite amplicon sequencing: Five different cell lines and DNA from normal colon mucosa were tested for their *AREG* and *EREG* promoter methylation. Per sequenced sample, 40 to 50 reads are shown, represented by the lines. CpG-positions are represented by the columns. Genomic positions are defined by the first proximal and last distal CpG of the analysed amplicon. Averaged DNA methylation and total number of reads are given at the right of the respective map. Blue: the CpG is unmethylated, Red: the CpG is methylated, White: the methylation status was not detected.

AREG and *EREG* promoters were mainly unmethylated in the cell lines CaCO2, HCT116, LIM1215 and SW480 and the colon mucosa. In RKO cells, the *AREG* promoter was also mainly unmethylated, however, the *EREG* promoter region was highly methylated.

Since *AREG* and *EREG* mRNA and protein levels were shown to be up-regulated in the cells by DAC, but the promoters were mainly unmethylated in the tested cells, the epigenetic regulatory regions might be at different regions within the genes. To identify these regions Dr. Sascha Tierling screened for CpG-positions within the *AREG* and *EREG* genes, which were methylated and lost their methylation among DAC treatment. He identified two CpGs in both genes by msSNuPE experiments fulfilling these criteria

(data not shown, patent filed). The positions within the genes are shown in figure 13. The CpGs within the *AREG* gene were named CpG p150 and CpG p220, the CpGs within the *EREG* gene were named CpG p143 and p297.

Figure 13:

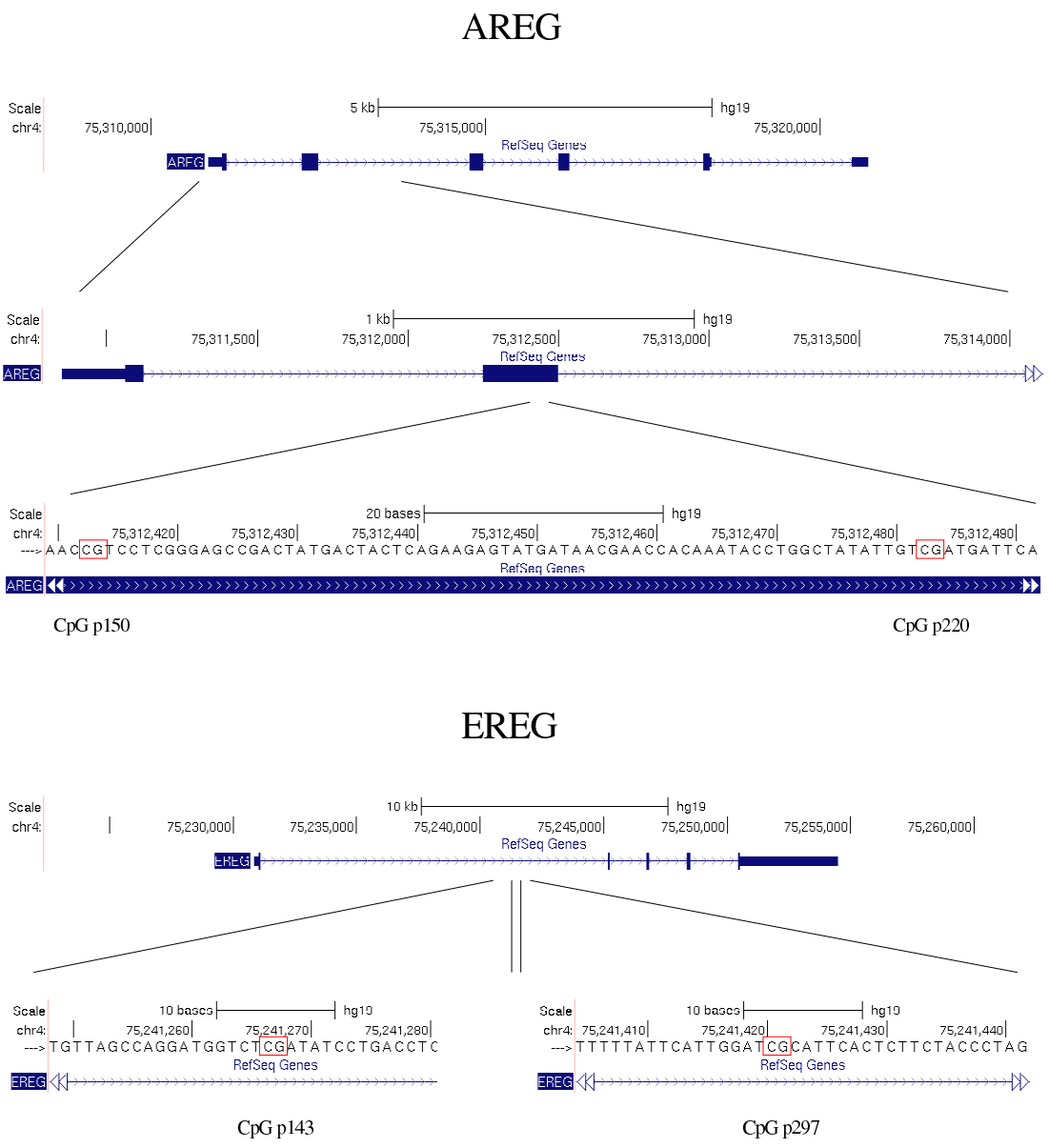


Figure 13: Chromosomal positions of *AREG* and *EREG* genes and intragenic CpGs: The intragenic *AREG* CpGs p150 and p220 are located within the *AREG* exon 2, whereas the intragenic *EREG* CpGs p143 and p297 are located within the *EREG* intron 1.

3.4 The methylation of intragenic CpG-sites within the *AREG* and *EREG* genes varies among different colorectal cancer cell lines

Figure 14:

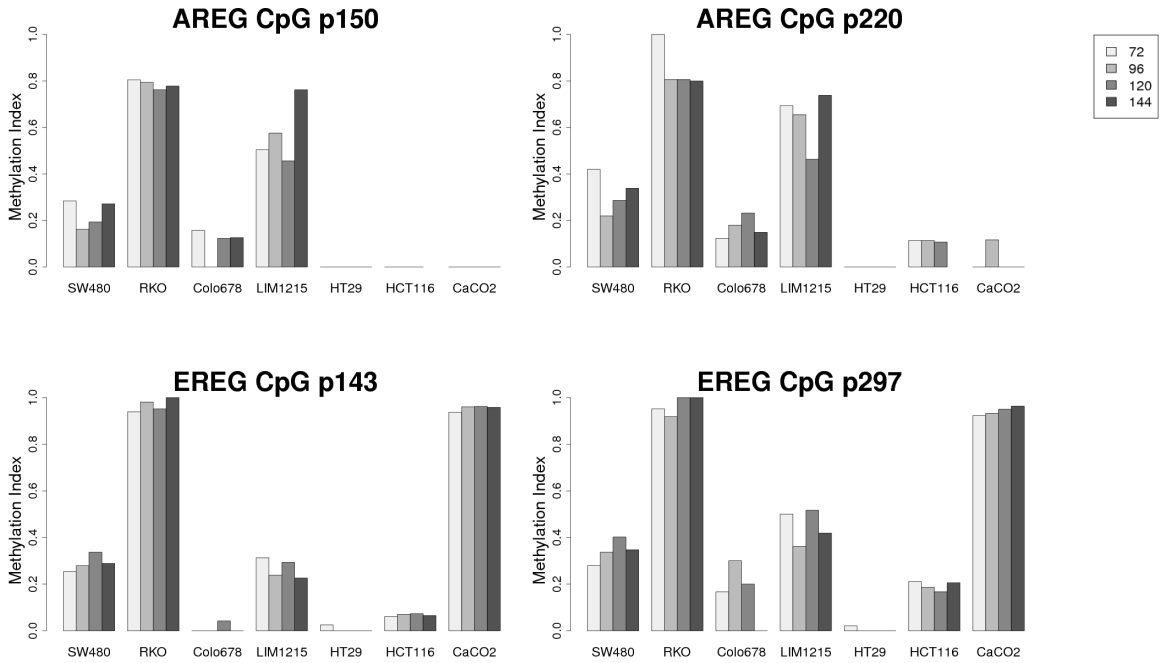


Figure 14: Methylation of *AREG* and *EREG* intragenic CpG-sites in different colorectal cancer cells: Seven colorectal cancer cells were tested for their intragenic methylation at *AREG* CpG p150 and p220 and at *EREG* CpG p143 and p297 (see figure 13) at four different timepoints. The methylation index shows the content of methylation indicating products within the primer extension products in msSNuPE-experiments and is an indicator for DNA methylation at the given CpG. A methylation index of 1 means fully methylated whereas a methylation index of 0 means fully unmethylated. Cells were sorted according to their *AREG* mRNA expression (see figure 3).

In section 3.1 was shown that the tested cell lines could be divided into low, medium and high AREG expressing cells. The comparison of figure 3 and figure 14 shows that the cells of the AREG high expressing group (HT29, HCT116 and CaCO2) were completely unmethylated at the *AREG* CpG p150 (MI = 0) and showed little methylation at *AREG* CpG p220 (Methylation index < 0.2). In contrast, methylation of these CpGs was observed among the AREG low and medium expressing cells. RKO and LIM1215 showed high methylation (MI > 0.5). Also SW480 showed a higher methylation than the cells of the AREG high expressing group (CpG p150: MI > 0.1 with an average of 0.23, CpG220: MI > 0.2 with an average of 0.32). With the exception of Colo678 (CpG p150: MI average of 0.1, CpG p220: MI average of 0.18), methylation of the two *AREG* CpGs seemed to correlate with the *AREG* mRNA expression.

However, *EREG* mRNA expression did not correlate with the methylation at the

identified CpGs within the *EREG* gene (see figure 3). Within the *EREG* medium and high mRNA expressing cells, HT29 was not methylated at the CpG p143 and CpG p297, whereas CaCO2 was fully methylated. Also in the *EREG* low expressing cells, the methylation index within RKO and Colo678 ranged from 1 to 0. Thus, methylation of CpGs p143 and p297 did not correlate with the *EREG* mRNA expression.

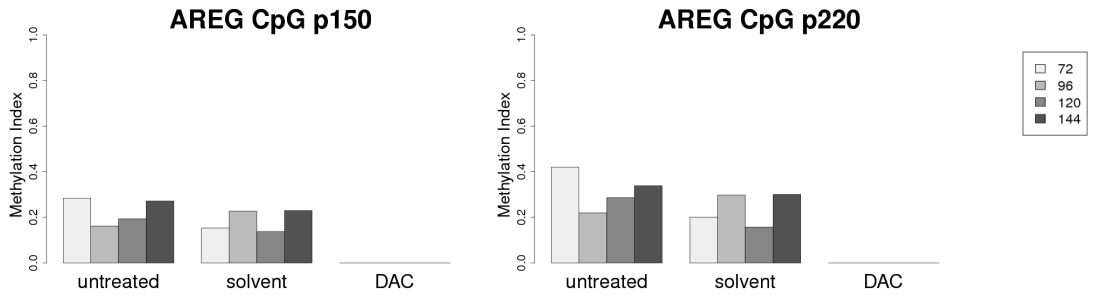
As a next step it was tested if DAC-treatment leads to a decrease of methylation in the observed CpGs to clarify if the increase of gene expression observed earlier could be explained by differential methylation within these regions.

3.5 Epigenetic compounds change the methylation of the intragenic CpG-sites in *AREG* and *EREG* genes

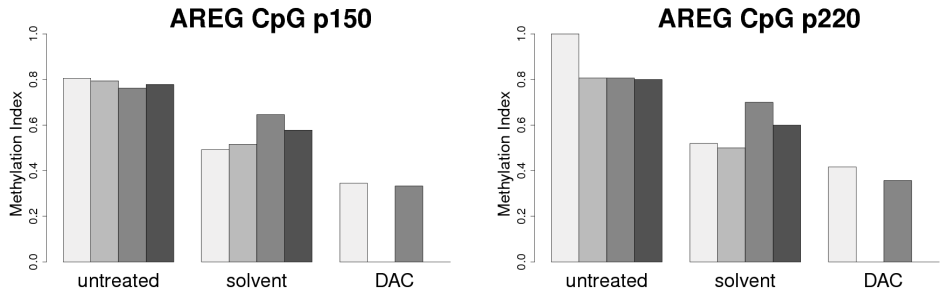
SW480, RKO and LIM1215, the cells exhibiting the highest methylation indices at *AREG* CpG p150 and CpG p220, were tested for methylation of these CpGs after treatment with the DNA methyltransferase inhibitor DAC (see figure 15). SW480 cells became fully unmethylated in both CpGs after 72 h of treatment. Demethylation remained up to 144 h of treatment. In LIM1215 and RKO cells, treatment with DAC also led to a decrease of methylation in both CpGs, however, methylation did not disappear completely in all timepoints as seen for SW480. Nevertheless, the methylation indices of both CpGs dropped in RKO cells from high ($0.8 < \mathrm{MI} < 1$) to low ($0 < \mathrm{MI} < 0.3$) and in a similar way in LIM1215 from high ($0.5 < \mathrm{MI} < 0.8$) to low ($0.1 < \mathrm{MI} < 0.3$). Since it was shown earlier that *AREG* mRNA expression correlated with methylation in *AREG* CpG p150 and CpG p220, an explanation of the *AREG* mRNA expression change after DAC treatment might be due to the demethylating effect of DAC on these CpGs.

Figure 15:

SW480:



RKO:



LIM1215:

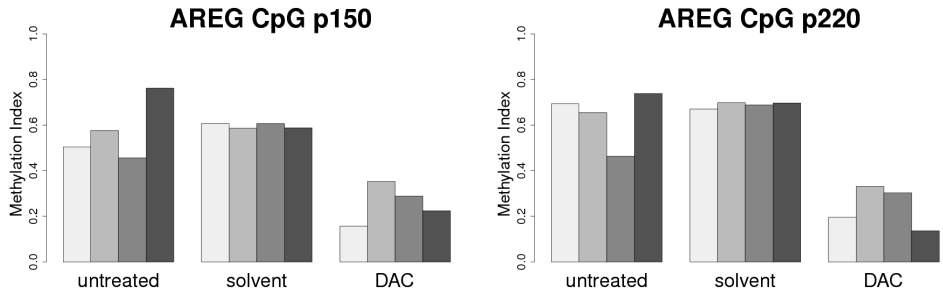
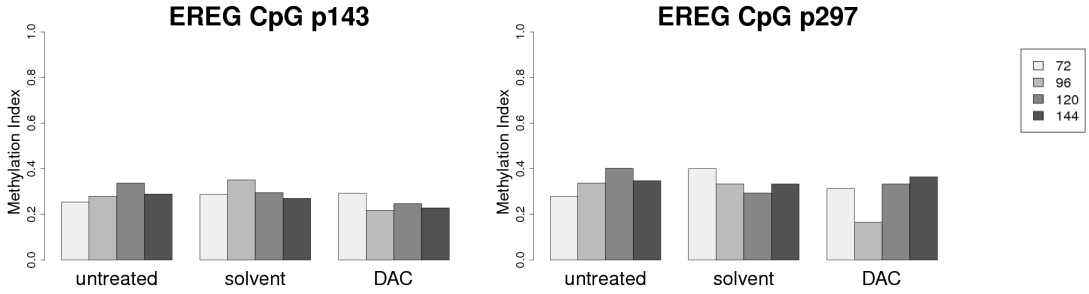


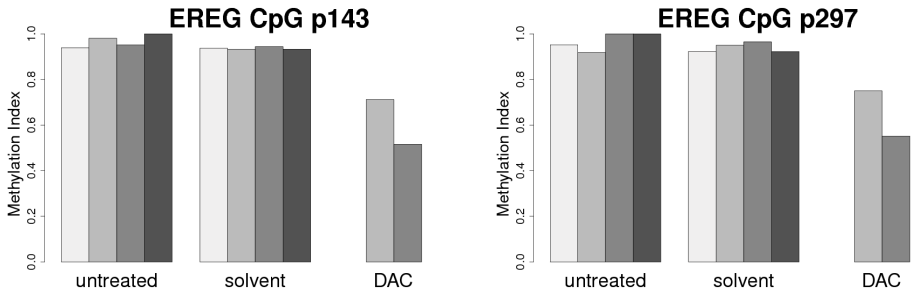
Figure 15: *AREG* intragenic methylation after DAC treatment: Three cancer cell lines were tested for their intragenic methylation at *AREG* CpGs p150 and p220 after treatment with 2.5 µM DAC at four different timepoints. The methylation index (MI) shows the content of methylation indicating products within the primer extension products in msSNuPE-experiments and is an indicator for DNA methylation at the given CpG. A methylation index of 1 means fully methylated whereas a methylation index of 0 means fully unmethylated.

Figure 16:

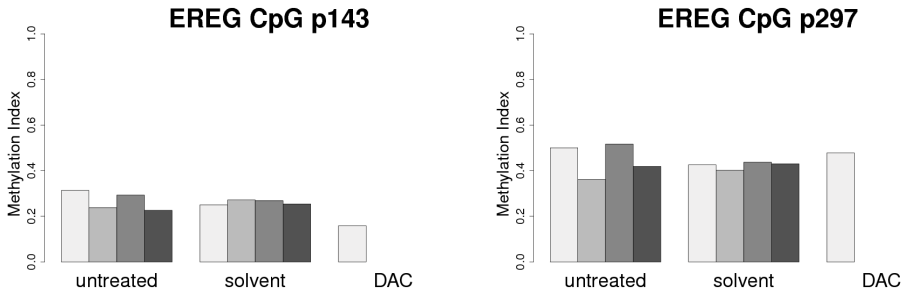
SW480:



RKO:



LIM1215:



CaCO2:

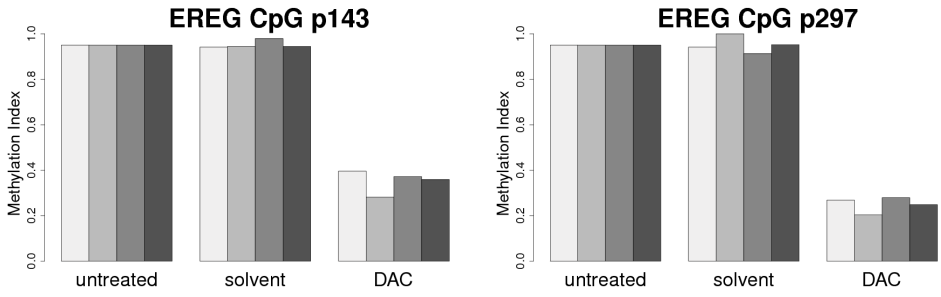
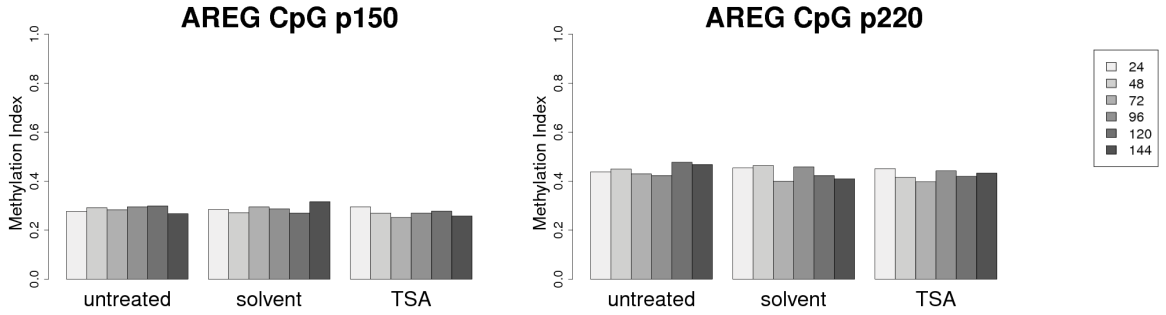


Figure 16: *EREG* intragenic methylation after DAC treatment: Four cancer cell lines were tested for their intragenic methylation at *EREG* CpGs p143 and p297 after treatment with 2.5 µM DAC at four different timepoints. The methylation index (MI) shows the content of methylation indicating products within the primer extension products in msSNuPE-experiments and is an indicator for DNA methylation at the given CpG. A methylation index of 1 means fully methylated whereas a methylation index of 0 means fully unmethylated.

The intragenic *EREG* CpGs p143 and p297 were methylated in the four cell lines SW480, RKO, LIM1215 and CaCO2. Therefore, it was also tested if methylation changed after DAC treatment (see figure 16). In contrast to what was observed for *AREG*, methylation of the *EREG* CpGs did not decrease in SW480 and LIM1215, although SW480, for example, had a methylation index similar to the *AREG* CpGs. In RKO and CaCO2 methylation of the CpGs decreased. The strongest effect was observed in CaCO2. However, *EREG* mRNA expression did not change in CaCO2 after DAC treatment (see figure 4). Unfortunately, the samples RKO 72 h and 144 h as well as LIM1215 96 h - 144 h did not give a result at the msSNuPE experiment.

Figure 17:

SW480:



LIM1215:

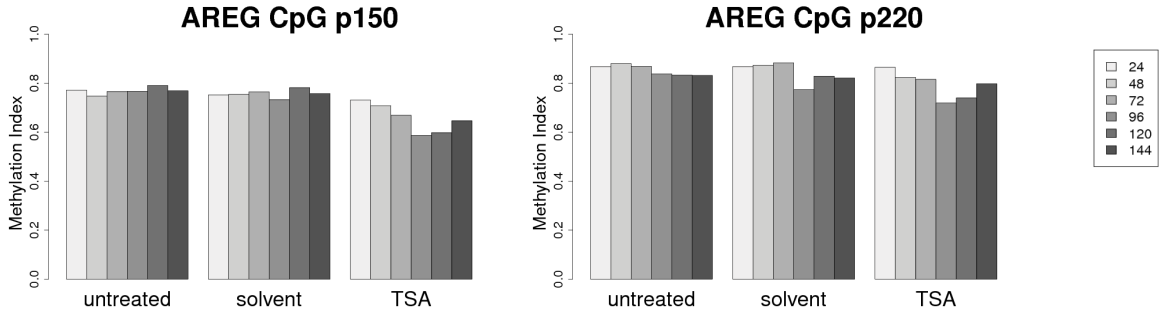


Figure 17: *AREG* intragenic methylation after TSA treatment: SW480 and LIM1215 were tested for their intragenic methylation at *AREG* CpG p150 and p220 after treatment with 25 ng/ml TSA at six different timepoints. The methylation index (MI) shows the content of methylation indicating products within the primer extension products in msSNuPE-experiments and is an indicator for DNA methylation at the given CpG. A methylation index of 1 means fully methylated whereas a methylation index of 0 means fully unmethylated.

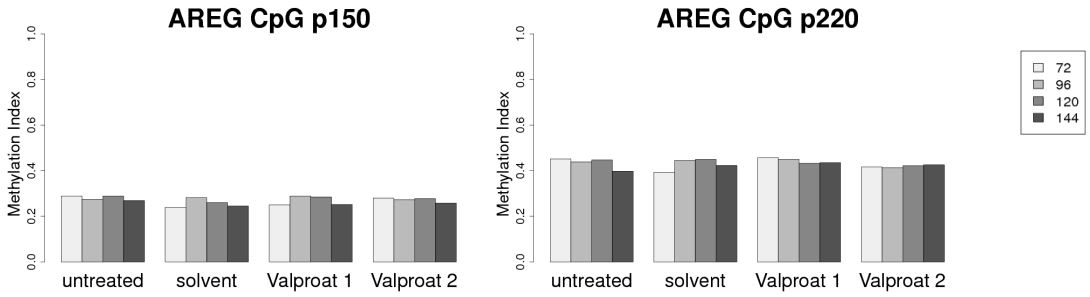
In addition to investigating the effect of DAC treatment it was also tested, whether HDACis have an effect on the intragenic methylation of the *AREG* gene, since treatment with HDACis mainly led to an increase of AREG expression, too (see section 3.2.2). The effect of TSA onto the methylation of *AREG* CpGs p150 and p220 is shown for the cells SW480 and LIM1215 in figure 17. Both cells differ in their initial methylation index, but in contrast to DAC treatment, methylation of the tested CpGs was changed upon

TSA treatment in none of the cells. However, in LIM1215 a reduction of CpG p150 methylation was observed from MI = 0.8 to MI = 0.6. A tendency was also seen for CpG p220 but to a lower extend.

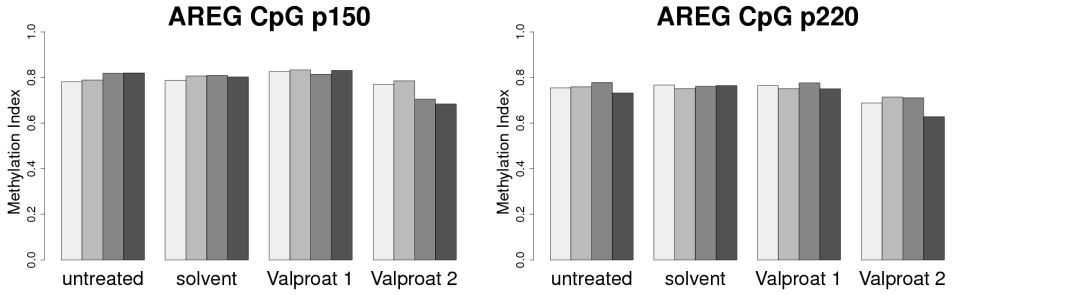
Besides the effect of TSA, the effect of Valproat on the methylation of the intragenic *AREG* CpGs (see figure 18) and *EREG* CpGs (see figure 19) was tested, too. Neither *AREG* nor *EREG* methylation was influenced by Valproat in the tested cells.

Figure 18:

SW480:



RKO:



LIM1215:

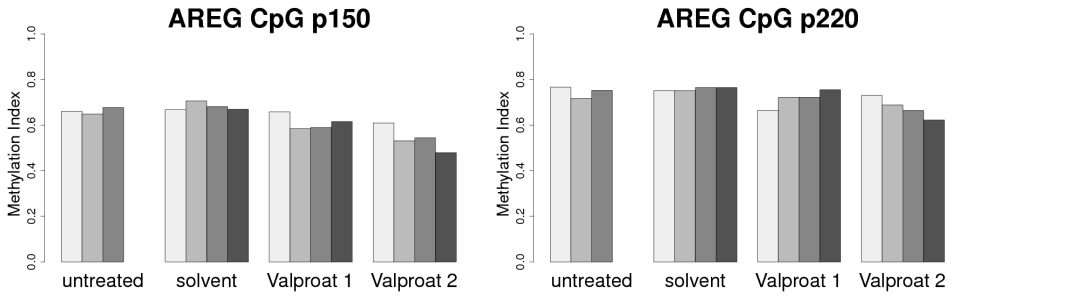
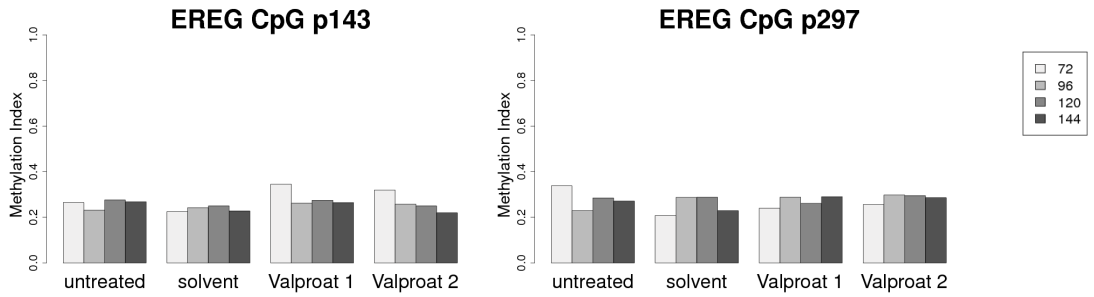


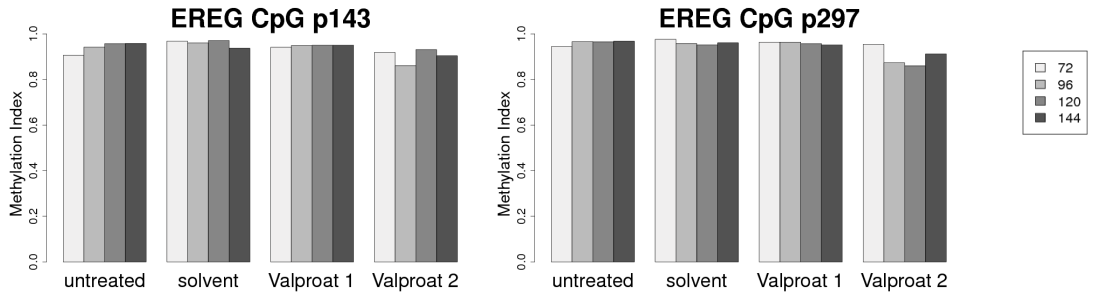
Figure 18: *AREG* intragenic methylation after Valproat treatment: SW480, RKO and LIM1215 were tested for their intragenic methylation at AREG CpG p150 and p220 after treatment with 1 mM or 2 mM Valproat (Valproat 1 and 2) in four different timepoints. The methylation index shows the content of methylation indicating products within the primer extension products in msSNuPE-experiments and is an indicator for DNA methylation at the given CpG. A methylation index of 1 means fully methylated whereas a methylation index of 0 means fully unmethylated.

Figure 19:

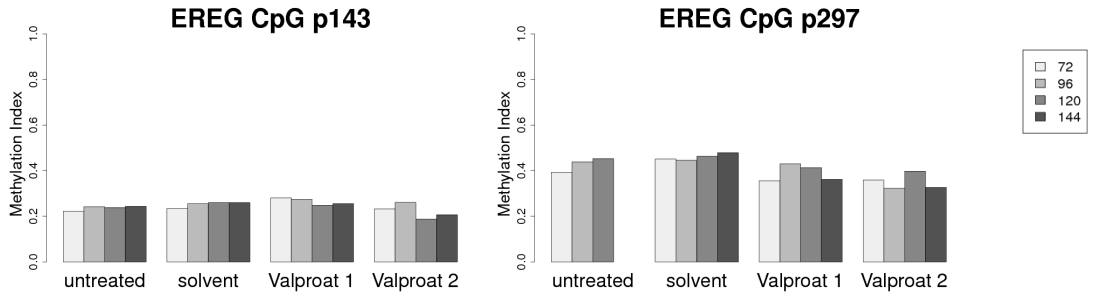
SW480:



RKO:



LIM1215:



CaCO2:

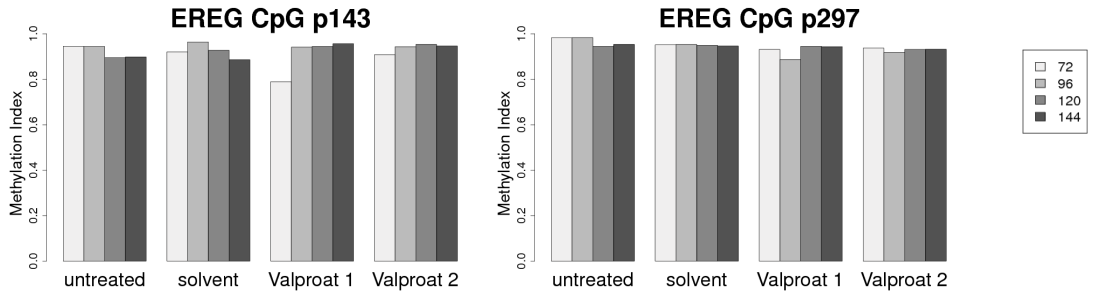


Figure 19: *EREG* intragenic methylation after Valproat treatment: Four different cell lines were tested for their intragenic methylation at EREG CpG p143 and p297 after treatment with 1 mM or 2 mM Valproat (Valproat 1 and 2) in four different timepoints. The Methylation index was calculated after msSNuPE experiment. A methylation index of 1 means fully methylated whereas a methylation index of 0 means fully unmethylated.

To summarize, DAC treatment led to a decrease of methylation at the intragenic *AREG* CpGs. Contrary, DAC had a lower effect on the methylation of the intragenic CpGs of the *EREG* gene. The HDACis TSA and Valproat did not influence methylation.

3.6 The *AREG* intragenic CpG-site has methylation-dependent promoter function

As shown in section 3.4, AREG expression correlated with the intragenic methylation at CpG p150 and p220 within *AREG* exon 2. To address the biological function of these CpGs, a promoter function of exon 2 was tested. For that, plasmids were generated with the CpG-free plasmid pCpGl-basic as backbone[239] (see figure 1). Because pCpGl-basic is a CpG-free plasmid, *in vitro* methylation of all plasmids would only lead to methylation within the insert, but not in the plasmid. Thus, a potential methylation-dependent promoter function can be analyzed.

Plasmid generation is described in section 2.8.1. Four groups of plasmids were made, which are described in figures 20 to 23. PCpGl-basic derived plasmids were used to analyze functions of *AREG* promoter and exon 2 directly (see figure 20). Besides the *AREG*-promoter, also the *EF1*-promoter and a Tata-binding motif were cloned into the pCpGl-basic plasmid to be combined in subsequent steps with the *AREG* exon 2 sequence. The obtained pCpGl-AREG-promoter derived plasmids (see figure 21) as well as pCpGl-EF1 and pCpGl-Tata derived plasmids (see figures 22 and 23) were used to analyze *AREG* exon 2 function, when coupled to a promoter or promoter element. PCpGl-AREG-exon2-R, pCpGl-AREG-exon2-Tata and pCpGl-AREG-exon2-R-Tata were methylated *in vitro* to analyze methylation dependency (see figure 24).
A CTCF-binding site located within *AREG* exon 2, was mutated in the plasmids pCpGl-AREG-exon2-R, pCpGl-AREG-exon2-AREG-promoter, pCpGl-AREG-exon2-R-AREG-promoter and pCpGl-AREG-exon2-R-Tata, to analyze the influence of this transcription factor binding site on promoter function (see figure 24). All pCpGl-basic derived plasmids contain a *Firefly-Luciferase* reporter-gene (Luc). Luminescence was used to evaluate promoter function, after supplying firefly luciferin to the extracts of the transfected cells. As a transfection control a *Renilla-Luciferase* reporter-gene containing plasmid (Ren) was used in each experiment (see section 2.8.3).

Figure 20:

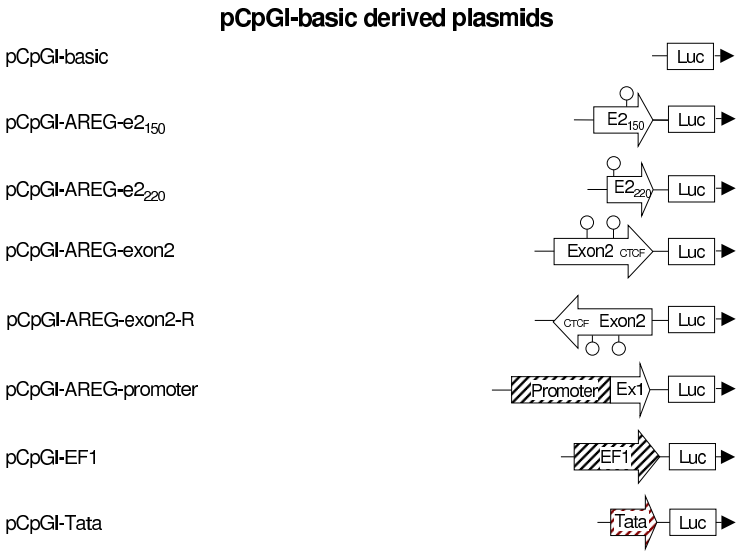


Figure 20: pCpGl-basic derived plasmids for promoter function analysis: The complete sequence and two parts of the sequence of *AREG* exon 2 were cloned into the plasmid pCpGl-basic (see figure 1). The intragenic CpGs p150 and p220 are located within exon 2 of the *AREG* gene (represented by white circles). A CTCF-binding site is located within exon 2 of the *AREG* gene, too (represented by "CTCF"). Also the sequence of the *AREG*-promoter, the *EF1*-promoter and the Tata-binding motif were cloned into the plasmid (see section 2.8.1). The insert-orientations are indicated (arrows) in front of the *Firefly-Luciferase* gene (Luc). The name of each plasmid is shown how they are used in the text.

Figure 21:

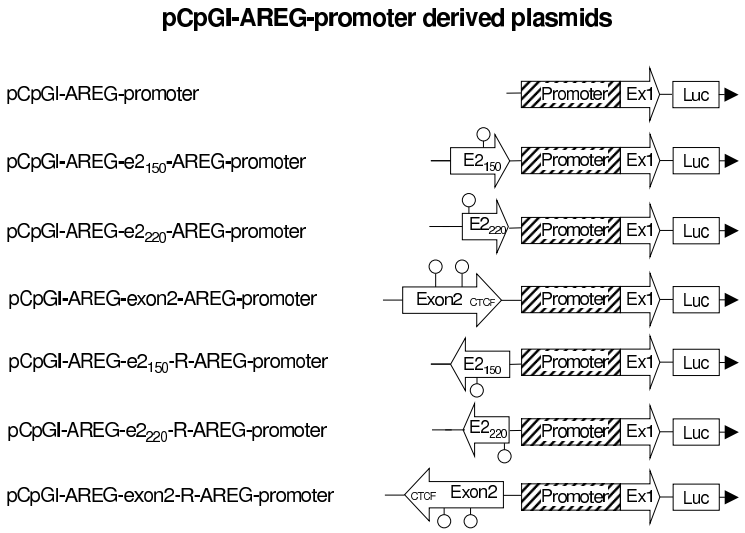


Figure 21: pCpGl-AREG-promoter derived plasmids for promoter function analysis: The complete sequence and two parts of the sequence of *AREG* exon 2 were cloned into the plasmid pCpGl-AREG-promoter (see figure 20) The intragenic CpGs p150 and p220 are represented by white circles, a CTCF-binding site is indicated by "CTCF". The insert-orientations are indicated (arrows). The name of each plasmid is shown how they are used in the text.

Figure 22:

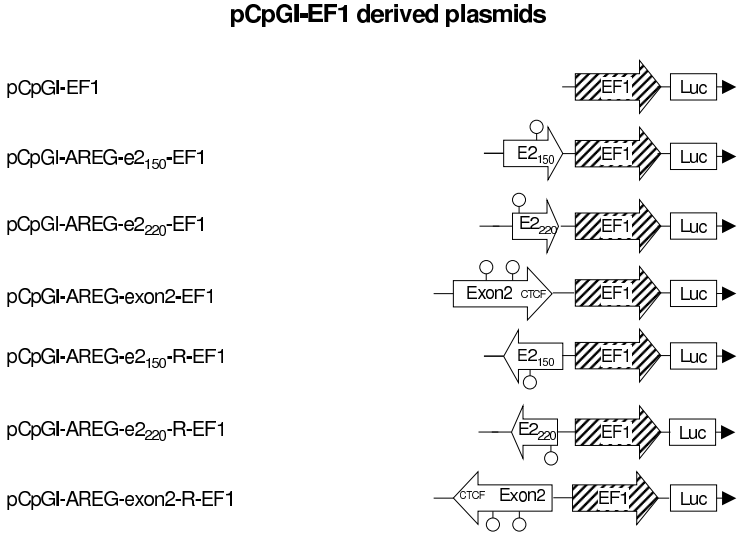


Figure 22: pCpGl-EF1 derived plasmids for promoter function analysis: The complete sequence and two parts of the sequence of *AREG* exon 2 were cloned into the plasmid pCpGl-EF1 (see figure 20) The intragenic CpGs p150 and p220 are represented by white circles, a CTCF-binding site is indicated by "CTCF". The insert-orientations are indicated (arrows). The name of each plasmid is shown how they are used in the text.

Figure 23:

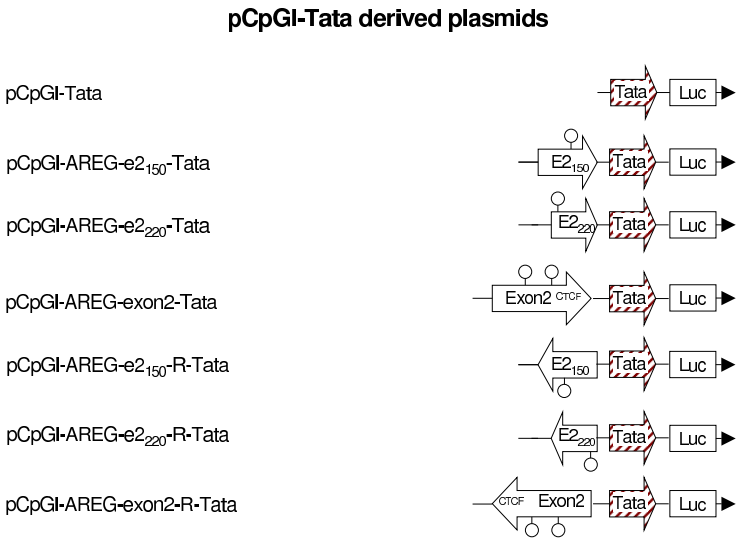


Figure 23: pCpGl-Tata derived plasmids for promoter function analysis: The complete sequence and two parts of the sequence of *AREG* exon 2 were cloned into the plasmid pCpGl-Tata (see figure 20) The intragenic CpGs p150 and p220 are represented by white circles, a CTCF-binding site is indicated by "CTCF". The insert-orientations are indicated (arrows). The name of each plasmid is shown how they are used in the text.

Figure 24:

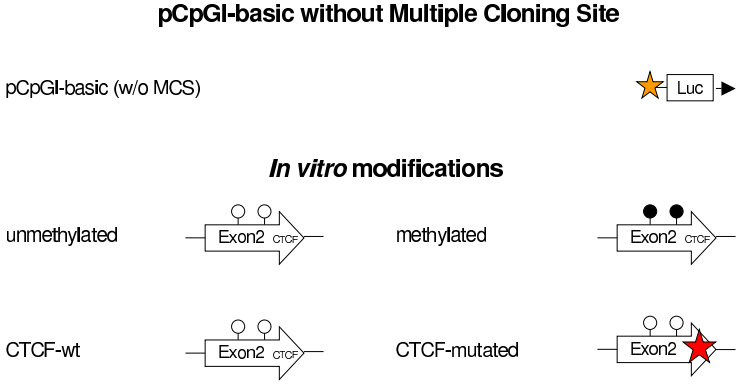


Figure 24: pCpGl-basic w/o MCS and *in vitro* modifications of the plasmids for promoter function analysis: To address the effect of the MCS on the promoter function, a MCS-free pCpGl-basic was created. To address methylation-specific effects on the promoter function and the effect of the CTCF-binding site located within *AREG* exon 2, plasmids were modified by methylation (represented by black circles) or by mutation (represented by red star).

In LIM1215 cells, an increased promoter function was observed for the *AREG* promoter-containing plasmid (see figure 25 A). *AREG* exon 2 and also the tested *AREG* exon 2 parts did not have promoter function within the tested plasmids. Also, no increase in promoter function was seen for the CpG-free EF1-promoter and the Tata-binding motif. Next, the sequences of *AREG* exon 2 or the *AREG* exon 2 parts $e2_{150}$ and $e2_{220}$ were coupled to the *AREG* promoter-sequence containing pCpGl-basic derived plasmid (see figure 25 B). The sequences of $e2_{150}$ as well as $e2_{220}$ did not change promoter function of the *AREG* promoter neither in forward nor in reverse orientation. A minimal increase of promoter function was observed when coupling the *AREG* exon 2 sequence completely to the promoter. But this increase was not significant. The promoter function analysis of the *AREG* exon 2 coupled to the EF1 is shown in figure 25 C. Compared to EF1, none of the exon 2 derived sequences increased promoter function. Actually, all observed values are near to zero.

When coupling *AREG* exon 2 in reverse orientation to the Tata protein binding motif (see figure 25 D), promoter function was significantly increased compared to the Tata-binding motif alone. But when coupling any other sequence to the Tata-binding motif, no significant increase was observed. However, as seen in figure 25 D, the experiment was accompanied with very high standard deviations of the measurements, which are caused, for example, by bad transfection efficiency. As a consequence, the results are not trustworthy for the very low measured data. Therefore, LIM1215 cells are substituted by HCT116 cells and CaCO2 cells in the following experiments.

Figure 25:

LIM1215:

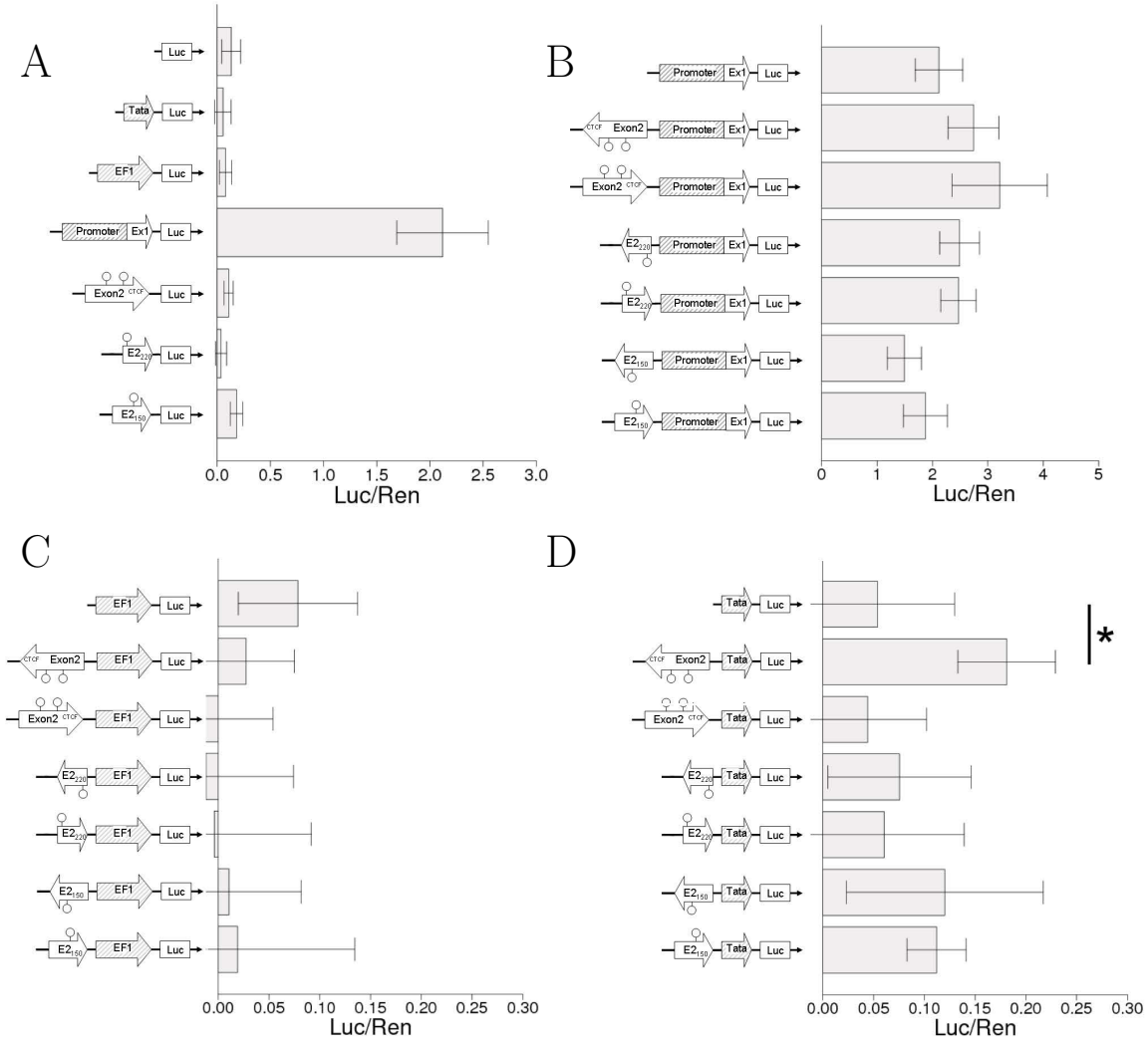


Figure 25: Promoter function analysis in LIM1215 cells: Promoter analysis was performed by luminescence measurement using the DLR Assay (Promega) after transfection into LIM1215. Firefly-luciferase derived luminescence (Luc) was normalized by Renilla-luciferase derived luminescence (Ren). A: pCpGl-basic derived plasmids, B: pCpGl-AREG-promoter derived plasmids, C: pCpGl-EF1 derived plasmids, D: pCpGl-Tata derived plasmids. A two-tailed heteroscedastic student's t-test was performed for some of the samples: *: p-value < 0.05 , **: p-value < 0.01

Figure 26:

HCT116:

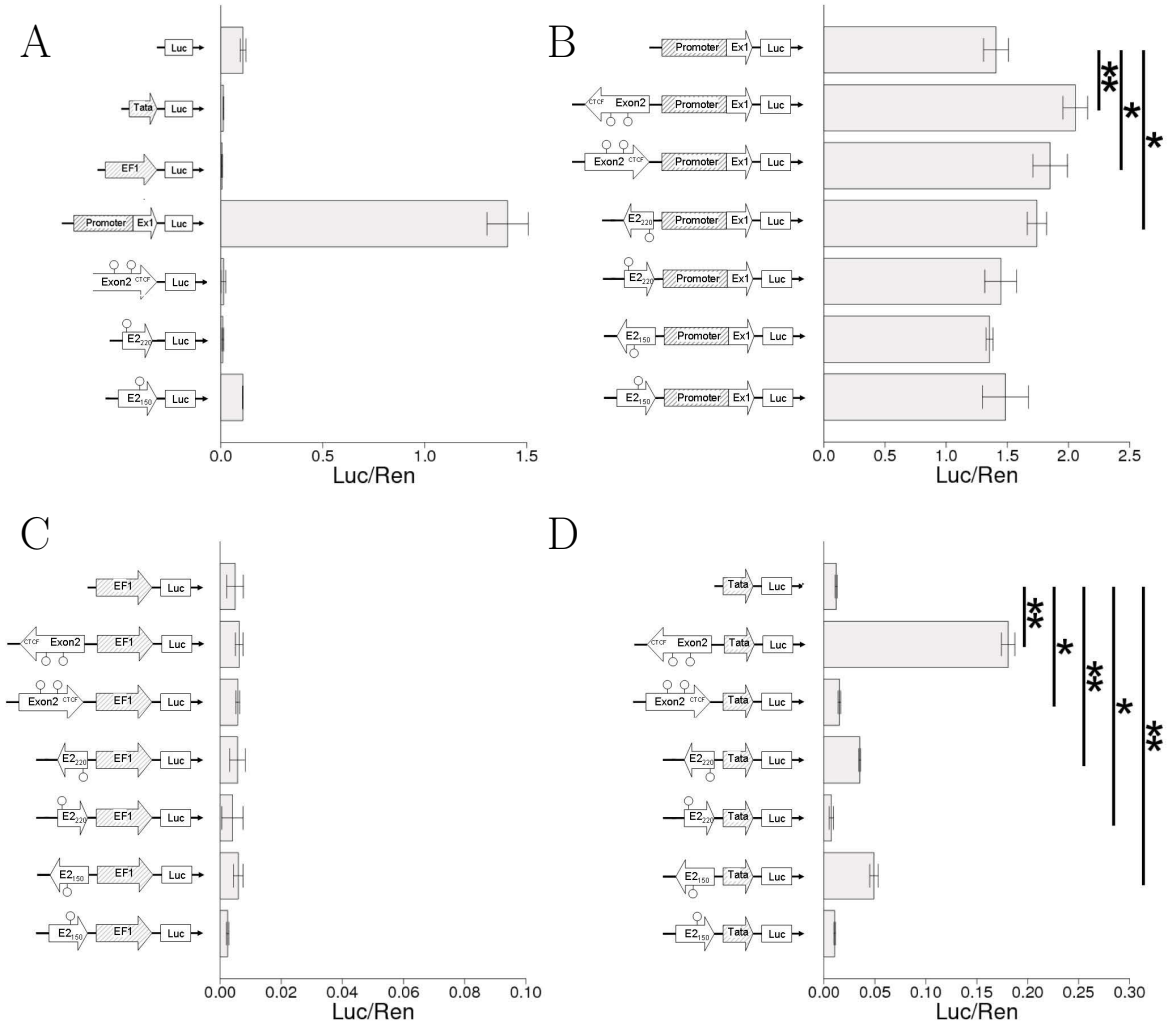


Figure 26: Promoter function analysis in HCT116 cells: Promoter analysis was performed by luminescence measurement using the DLR Assay (Promega) after transfection into HCT116. Firefly-luciferase derived luminescence (Luc) was normalized by Renilla-luciferase derived luminescence (Ren). A: pCpGl-basic derived plasmids, B: pCpGl-AREG-promoter derived plasmids, C: pCpGl-EF1 derived plasmids, D: pCpGl-Tata derived plasmids. A two-tailed heteroscedastic student's t-test was performed for some of the samples: *: p-value < 0.05 , **: p-value < 0.01

Similar to LIM1215, promoter function was also detected in HCT116 for the *AREG* promoter (figure 26 A). As before, no promoter function was seen for the EF1 promoter, the Tata-binding motif as well as the *AREG* exon 2 derived sequences. Besides the *AREG* exon 2 part $e2_{150}$ containing plasmid, the promoter functions of the measured plasmids were even lower than the function of the insert free pCpGl-basic.

When introducing the *AREG* exon 2 sequence into the *AREG* promoter sequence containing pCpGl-plasmid (figure 26 B), the promoter function increased significantly compared to the *AREG* promoter alone. Surprisingly, the promoter function increased stronger, when coupling the *AREG* exon 2 sequence in reverse orientation to the *AREG* promoter sequence. The increase was highly significant. Due to this result it is probable that the *AREG* exon 2 sequence rendered enhancer effects to the *AREG* promoter in HCT116 cells. The *AREG* exon 2 part $e2_{220}$ showed also a significant increase of promoter function, when coupled to the promoter in reverse orientation. However, the effect was lower than the effect of the complete *AREG* exon 2 sequence indicating that both CpGs, or the whole *AREG* exon 2 is necessary for the enhancer effect.

Similar to the results in LIM1215, EF1 as well as the EF1 derived plasmids did not have promoter function in HCT116 (see figure 26 C). Therefore, EF1 and EF1-derived plasmids were omitted from future experiments.

The Tata-binding motif did not show significant promoter function in HCT116 cells (figure 26 D). When introducing the *AREG* exon 2 sequence to the Tata-binding motif in forward orientation, promoter function increased. Interestingly, when introducing the *AREG* exon 2 sequence in reverse orientation, promoter function increased highly significant about 10-fold. Also the *AREG* exon 2 sequence parts, $e2_{150}$ and $e2_{220}$, increased promoter function 3-fold and 4-fold respectively when coupled to the Tata-binding motif in reverse orientation.

In the last two experiments it is obvious that the values obtained by the promoter assays were very low. To test whether the values were above background, an experiment was performed using HCT116 cells and the plasmids pCpGl-AREG-exon2-R-Tata and pCpGl-AREG-promoter, which differ in the values created by promoter function experiments by approximately one magnitude. After transfection, the cell extracts were diluted 7 times in 2-fold steps to determine the measurement background of the Luminoscan RS luminometer. The results are shown in figure 27.

Figure 27:

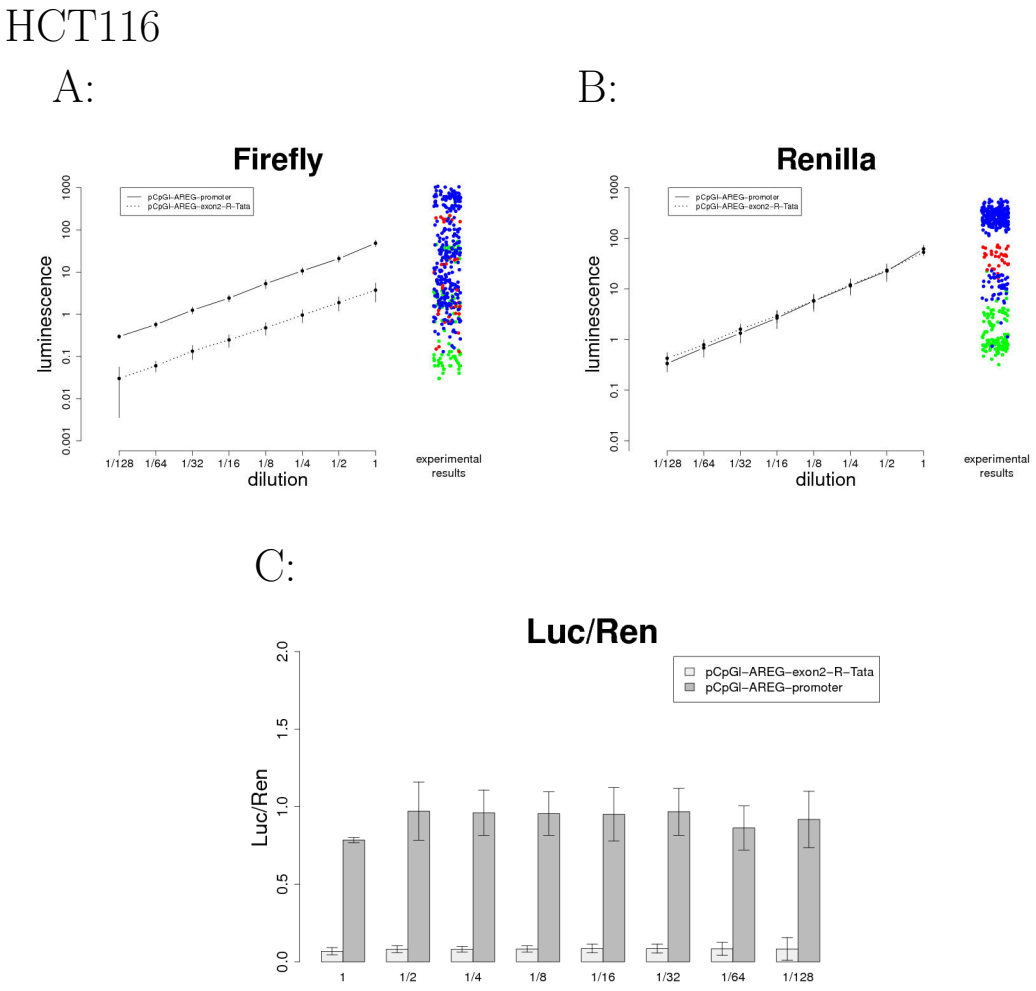


Figure 27: Validation of promoter function experiments: HCT116 cells were transfected with pCpGl-AREG-exon2-R-Tata or pCpGl-AREG-promoter together with the Renilla control plasmid. The cell lysates were diluted in 2-fold steps and luminescence was measured using the DLR Assay (Promega). A: Firefly-luciferase derived luminescence, B: Renilla-luciferase derived luminescence, A-B: "experimental results": values obtained in all promoter function experiments within this study: blue: HCT116, green: LIM1215, red: CaCO2. C: Firefly-luciferase derived luminescence (Luc) was normalized by Renilla-luciferase derived luminescence (Ren).

Firefly-luciferase derived luminescence as well as Renilla-luciferase derived luminescence was observed in all diluted samples (see figures 27 A and B). Besides the Firefly-luciferase derived luminescence of the 1/128 dilution, which had a very high error bar, the values showed a linear correlation to the dilution. The graphs describing this linear correlation did not flatten to a constant luminescence value, which means that the background of the Luminoscan RS luminometer was not reached, yet. Also, when normalizing the Firefly-luciferase derived luminescence by Renilla-luciferase derived luminescence, all dilutions gave similar results (see figure 27 C). The normalization procedure is applicable for values down to approximately 0.4 for Renilla- and 0.06 for Firefly-derived luminescence. As seen at the right sides of the Firefly and Renilla-plots

(see figures 27 A and B), all values within these study obtained for HCT116 and CaCO2 transfected cells lie above these thresholds. Only some values of the LIM1215 transfected cells lie below the value of the Firefly-luciferase derived luminescence, which might be due to the low transfection efficiency in LIM1215 cells mentioned earlier. To summarize, this experiment demonstrated that the values obtained in this study are valid and could be analyzed.

The results of the t-tests are a second issue when interpreting the results of the promoter function experiments. Although the p-values calculated, showed that the results of two plasmids differ significantly, the significance is sometimes hardly observable in the figures, especially when other plasmids led to much stronger results (see e.g. figure 26 D or figure 29 C). Nevertheless, to verify the t-tests, figure 26 D was enlarged as an example to show that the significance calculated is also visible in the plots (see supplementary material section 8.4).

In the following two experiments, it was analyzed, how methylation of *AREG* CpG p150 and p220 affects promoter function.

Figure 28:

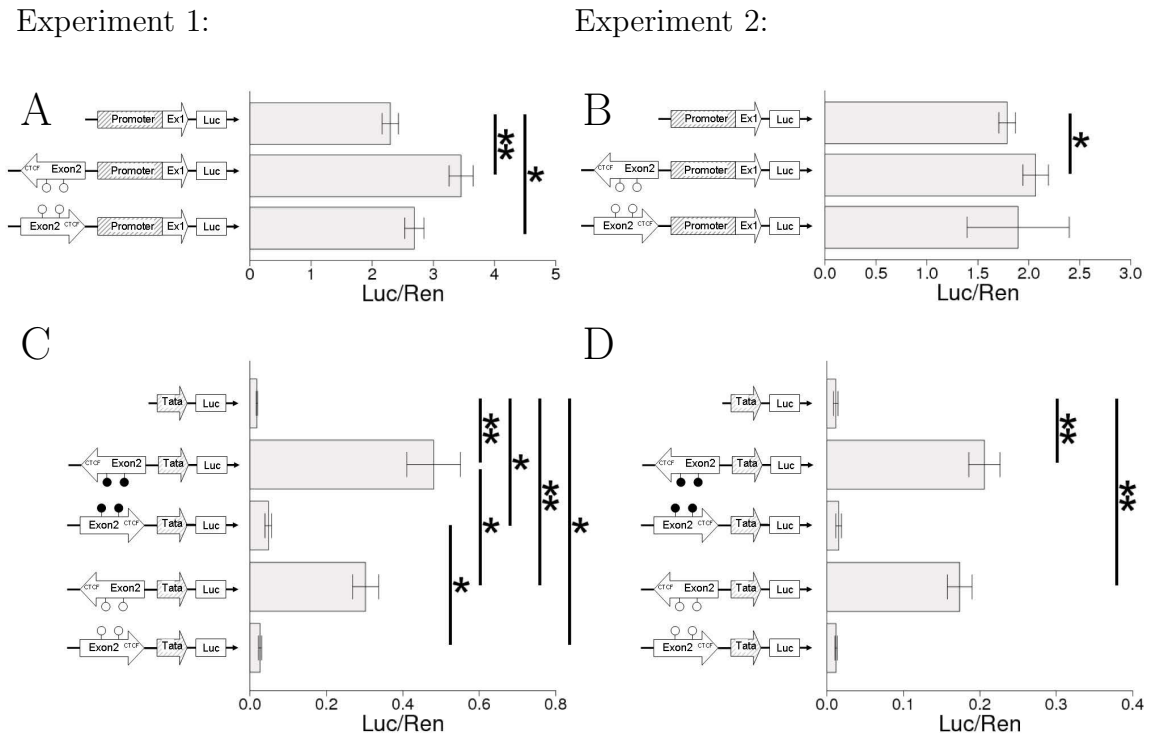


Figure 28: Methylation-dependent promoter function analysis in HCT116 cells: Promoter analysis was performed by luminescence measurement using the DLR Assay (Promega) after transfection into HCT116 cells. Firefly-luciferase derived luminescence (Luc) was normalized by Renilla-luciferase derived luminescence (Ren). Two independent experiments were made. (left and right). upper A and B: pCpGl-AREG-promoter derived plasmids, C and D: pCpGl-Tata derived plasmids. Black filled circles indicate an *in vitro* methylated insert. A two-tailed heteroscedastic student's t-test was performed for some of the samples: *: p-value < 0.05 , **: p-value < 0.01

In both additional experiments, the sequence of *AREG* exon 2 increased promoter function significantly when coupled in reverse orientation to the *AREG* promoter (see figure 28 A and B). Similar to the experiment before (see figure 26 B), the *AREG* exon 2 sequence also increased promoter function when coupled to the *AREG* promoter sequence in forward orientation (see figure 28 A).

When coupling the sequence of *AREG* exon 2 in reverse orientation to the Tata-binding motif (see figure 28 C and D), a highly significant increase of *AREG* promoter function was observed in both experiments and confirmed the earlier results. Interestingly, when methylating the sequence of *AREG* exon 2, the promoter function increased stronger when comparing with the unmethylated sequence. The effect was even significant in experiment 1. In the same experiment there was also observed a significant increase, when coupling *AREG* exon 2 in forward orientation to the Tata-binding motif. Similar to the reverse-oriented exon 2 sequence, a further significant increase was observed, when methylating the forward oriented AREG exon 2 sequence.

Figure 29:

CaCO2:

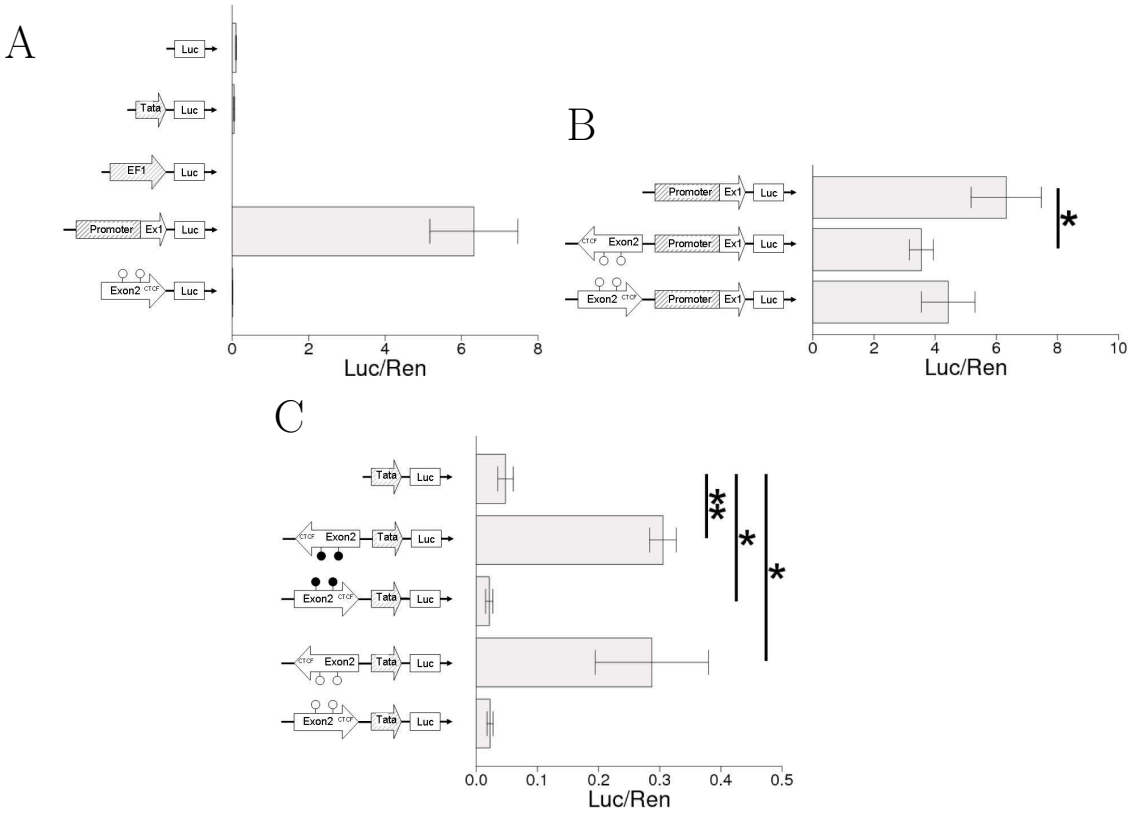


Figure 29: methylation-dependent promoter function analysis in CaCO2 cells: Promoter analysis was performed by luminescence measurement using the DLR Assay (Promega) after transfection into CaCO2 cells. Firefly-luciferase derived luminescence (Luc) was normalized by Renilla-luciferase derived luminescence (Ren). A: pCpGl-basic derived plasmids, B: pCpGl-AREG-promoter derived plasmids, C: pCpGl-Tata derived plasmids. Black filled circles indicate an *in vitro* methylated insert. A two-tailed heteroscedastic student's t-test was performed for some of the samples: *: p-value < 0.05 , **: p-value < 0.01

Promoter function of the plasmids was also analyzed in CaCO2 cells. (see figure 29). Similar to the other cell lines, *AREG* promoter showed promoter function whereas EF1 and the Tata-binding motif alone did not. (figure 29 A). In contrast to the other cell lines, coupling *AREG* exon 2 to the *AREG* promoter led to a decreased promoter function compared to the promoter alone (see figure 29 B and data not shown). This effect was orientation-independent. When coupling *AREG* exon 2 in forward orientation to the Tata-binding motif (figure 29 C), the promoter function of the Tata-binding motif decreased. The decrease was even significant for the methylated sequence. But similar to HCT116, coupling *AREG* exon 2 to the Tata-binding motif in reverse orientation led to a 6-fold increase of promoter function. However, in contrast to HCT116, additional methylation did not further change promoter function.

Figure 30:

HCT116:

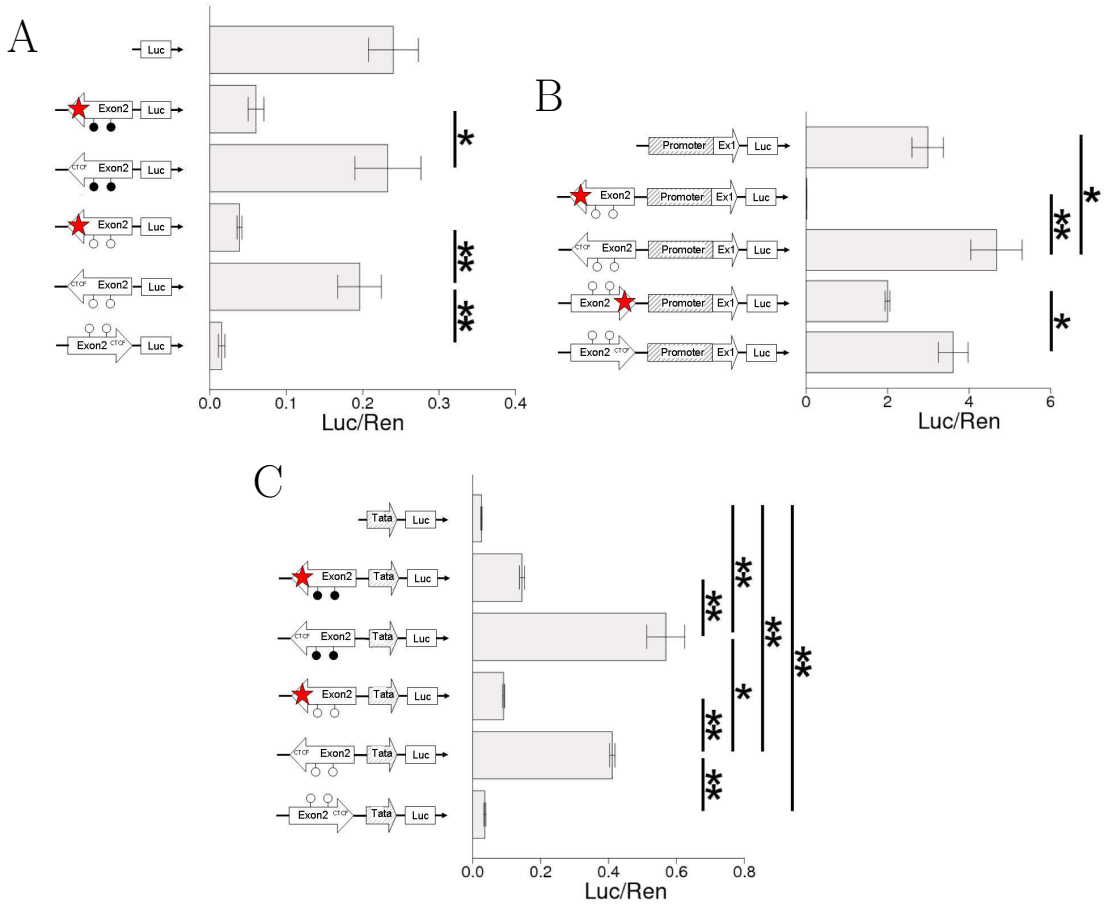


Figure 30: Methylation and CTCF dependent promoter function analysis in HCT116 cells: Promoter analysis was performed by luminescence measurement using the DLR Assay (Promega) after transfection into HCT116 cells. Firefly-luciferase derived luminescence (Luc) was normalized by Renilla-luciferase derived luminescence (Ren). A: pCpGl-basic derived plasmids, B: pCpGl-AREG-promoter derived plasmids, C: pCpGl-Tata derived plasmids. Black filled circles indicate an *in vitro* methylated insert, A red star indicates that the CTCF-binding site was mutated to a HindIII binding site. A two-tailed heteroscedastic student's t-test was performed for some of the samples: *: p-value < 0.05 , **: p-value < 0.01

To address the potential influence of the CTCF-binding site within the *AREG* exon 2 sequence on the promoter function, mutated plasmids were generated using the Phusion site-directed mutagenesis kit (Thermo Scientific). Afterwards, the promoter function was compared to the non-mutated counterparts. HCT116 was used for analysis, because it showed the best results in the experiments before (see figure 30).

The sequence of *AREG* exon 2 in reverse orientation had a highly significant stronger promoter function than the same sequence in forward orientation (see figure 30 A). But when mutating the CTCF-binding site within these sequences, the promoter function dropped significantly. When methylating the *AREG* exon 2 reverse-oriented sequence, promoter function increased compared to the unmethylated sequence, but

not significantly. But, similar to the unmethylated state, mutation of the methylated reverse-oriented *AREG* exon 2 sequence led to a significant drop in promoter function, too. Surprisingly, pCpGl-basic alone showed promoter function as high as the plasmid containing the *AREG* exon 2 reverse-oriented sequence.
When coupling the reverse-oriented *AREG* exon 2 sequence to the *AREG* promoter sequence, promoter function increased significantly similar to the experiments before (see figure 30 B). However, mutation of this plasmid led to a complete loss of promoter function. Also, similar to the experiments before an increase of promoter function of the AREG promoter was observed when coupling the *AREG* exon 2 sequence in forward orientation to the *AREG* promoter. However, the effect was not significant. But when mutating the CTCF region, promoter function decreased significantly.
Last, the promoter function of the *AREG* exon 2 sequence coupled to the Tata-binding motif was evaluated (see figure 30 C). As in the experiments before, the reverse-oriented, but not the forward oriented *AREG* exon 2 sequence increased promoter function of the Tata-binding motif highly significant. Also as seen before, methylation of the reverse-oriented *AREG* exon 2 increased promoter function significantly. And similar to the *AREG* promoter coupled plasmids, mutation of the CTCF-binding region in unmethylated as well as methylated reverse-oriented *AREG* exon 2 sequences led to a highly significant drop of promoter function in Tata-motif coupled plasmids.

In the final experiment of this study, the conflict was addressed that pCpGl-basic and pCpGl-AREG-exon2 showed similar promoter function. A new plasmid was generated which lacks the multiple cloning site. Also, the influence of the CTCF-binding site on promoter function was analyzed again. The results are shown in figure 31. The promoter function dropped significantly for the pCpGl-basic plasmid without MCS which indicates that the MCS has promoter function (see figure 31 A). Similar to the experiments before, mutating the CTCF-binding site of *AREG* exon 2 leads to a significant drop of promoter function compared to the unmutated plasmids when coupled to the *AREG* promoter sequence or the Tata-binding motif (see figure 31 B and C).

Figure 31:

HCT116:

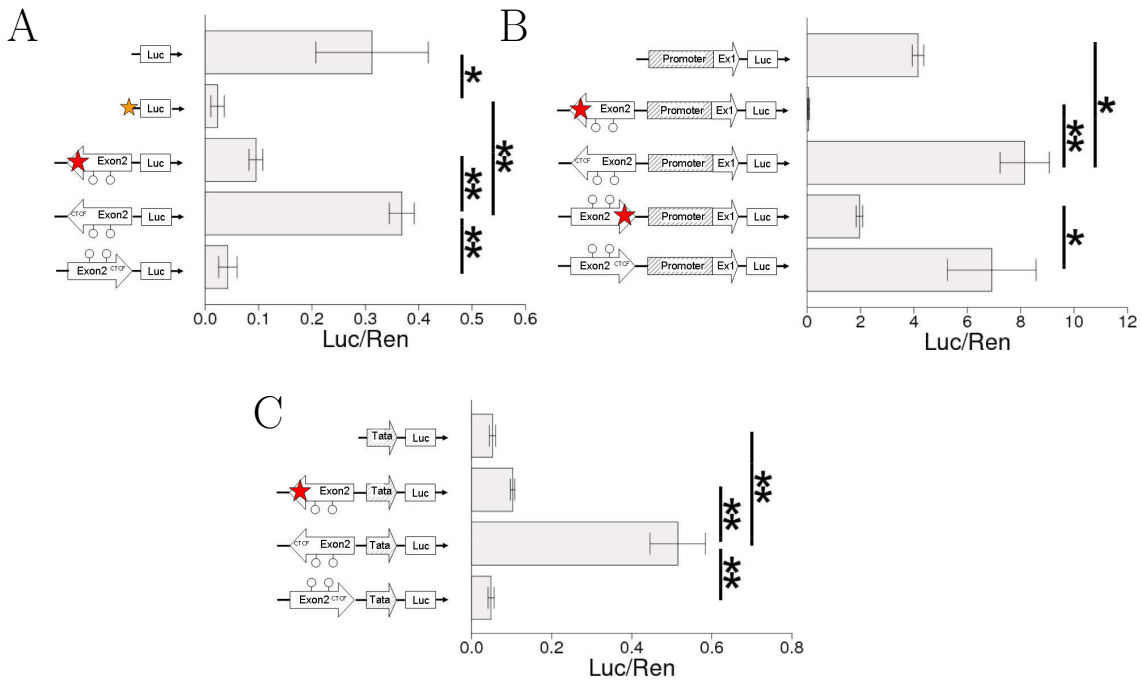


Figure 31: Repetition of CTCF-dependent promoter analysis and evaluation of the influence of the MCS: Promoter analysis was performed by luminescence measurement using the DLR Assay (Promega) after transfection into HCT116 cells. Firefly-luciferase derived luminescence (Luc) was normalized by Renilla-luciferase derived luminescence (Ren). A: pCpGl-basic derived plasmids, B: pCpGl-AREG-promoter derived plasmids, C: pCpGl-Tata derived plasmids. Black filled circles indicate an *in vitro* methylated insert, A red star indicates that the CTCF-binding site was mutated to a HindIII binding site. A two-tailed heteroscedastic student's t-test was performed for some of the samples a: *: p-value < 0.05 , **: p-value < 0.01

To summarize, the *AREG* exon 2 region, comprising CpG p150 and CpG p220 showed different promoter activity in different cells. In HCT116, it showed enhancer function in both orientations, when coupled to the *AREG* promoter. By mutating the CTCF site, the promoter function decreased strongly. When coupled to the Tata-binding motif in reverse orientation, the *AREG* exon 2 sequence highly increased promoter function. Interestingly methylation enhances this effect. Nevertheless, mutation of the CTCF-binding site had also a strong negative effect on promoter function in the methylated plasmids.

3.7 Zinc finger binding transcription factors are differentially expressed in colorectal cancer cell lines

Mutation of a CTCF-binding site within the *AREG* exon 2 sequence led to a drop of promoter function (see section 3.6). Therefore, the zinc finger binding transcription factors CTCF or CTCFL might contribute to *AREG* gene expression. To test this hypothesis, first *CTCF* as well as *CTCFL* mRNA levels were analyzed in different colorectal cancer cells. In addition, the mRNA expression of the zinc finger binding transcription factors *ZBTB4* and *ZBTB33* were also tested, because similar to CTCF and CTCFL, the proteins are able to bind CpG-containing DNA sequence motifs as well as CpG-free DNA sequence motifs.[242]

Figure 32:

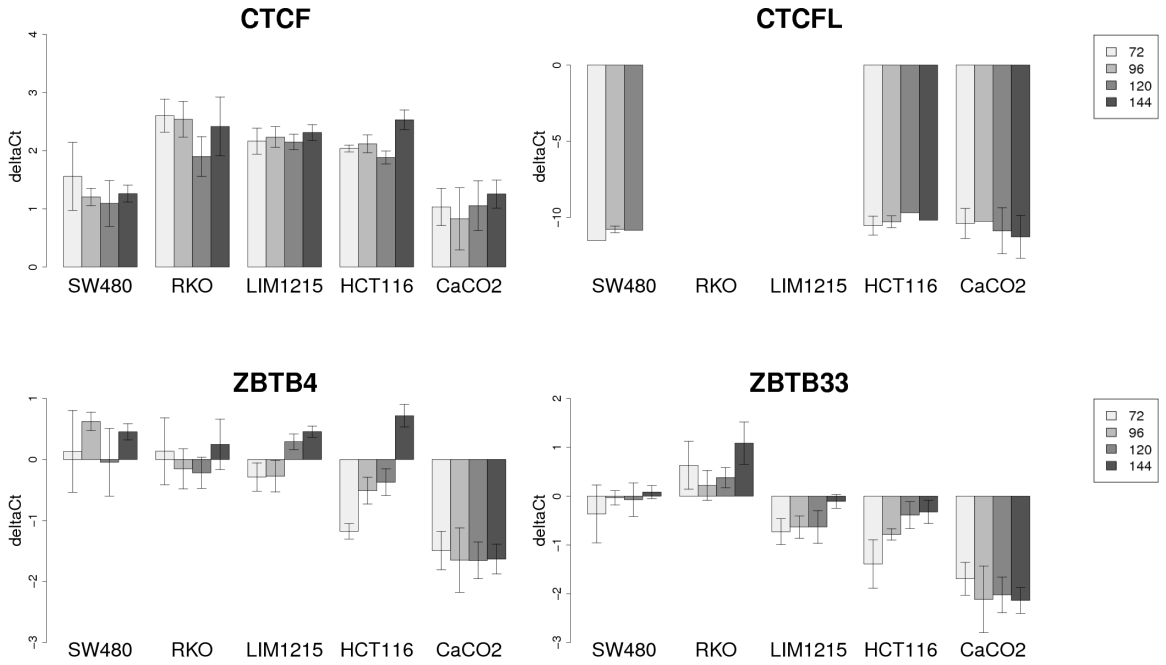


Figure 32: Transcription factor mRNA expression in in different colorectal cancer cells: Five colorectal cancer cell lines were tested for their *CTCF*, *CTCFL*, *ZBTB4* and *ZBTB33* mRNA expression from 72 h to 144 h after first medium change. Left: The ΔCt-values were calculated by subtraction of the *AREG* or *EREG* Ct-value from the control Ct-value. The cells were sorted according to their *AREG* mRNA expression (compare with figure 3).

In figure 32 the mRNA expression of *CTCF*, *CTCFL*, *ZBTB4* and *ZBTB33* is shown in five different colorectal cancer cell lines at four timepoints. The cancer cells were sorted according to their initial *AREG* gene expression (compare figure 3). *CTCF*-mRNA expression was present in all cell lines independent of their *AREG* expression, which is represented by high ΔCt-values. The AREG low expressing cell line SW480 as well as the AREG high expressing cell line CaCO2 showed a reduced mRNA expression compared to the remaining cell lines. *CTCFL* mRNA was not expressed in the cells. ΔCt-values of -10 and less indicate that the Ct-values for *CTCFL* rose at very

late cycle numbers, reflecting a very low mRNA level. RKO as well as LIM1215 showed no *CTCFL*-mRNA expression at all. Contrary to *CTCF*, *ZBTB4* as well as *ZBTB33* showed a differential mRNA-expression pattern in different cells. An inverse correlation between *AREG* mRNA expression and transcription factor expression was observed for *ZBTB33*. Low *AREG* mRNA expressing cells (SW480, RKO) expressed higher *ZBTB33* mRNA levels than high *AREG* mRNA expressing cells (HCT116, CaCO2). Also *ZBTB4* mRNA expression differed between the cells and similar to *ZBTB33*, CaCO2 showed the lowest mRNA expression. Furthermore, a time-dependent variation in all cells occurred except in CaCO2 showing that ZBTB4 and ZBTB33 might be regulated during proliferation. Although LIM1215 and HCT116 cells showed an increased *ZBTB4* and *ZBTB33* mRNA expression over time, the *AREG* mRNA expression remained stable (see figure 3). In HCT116 cells, the AREG protein level increased in the supernatant also time-dependently. However, this did not fit the inverse correlation seen for all cells. Nevertheless, other arguments strengthen the influence of ZBTB33 on the *AREG* expression.

Natalia Kuhn, a phd-student in the group of molecular tumorpathology at the Institute for Pathology (Charité Berlin), performed an experiment reducing the *ZBTB33* expression in LIM1215 cells by transfecting siRNAs targeting *ZBTB33* (see figure 33). She observed in 3 independent experiments that *AREG* and *EREG* mRNA expression is stronger in cells expressing less *ZBTB33*. In these experiments the inverse correlation between *AREG* and *ZBTB33* found in cell lines could be verified.

Figure 33:

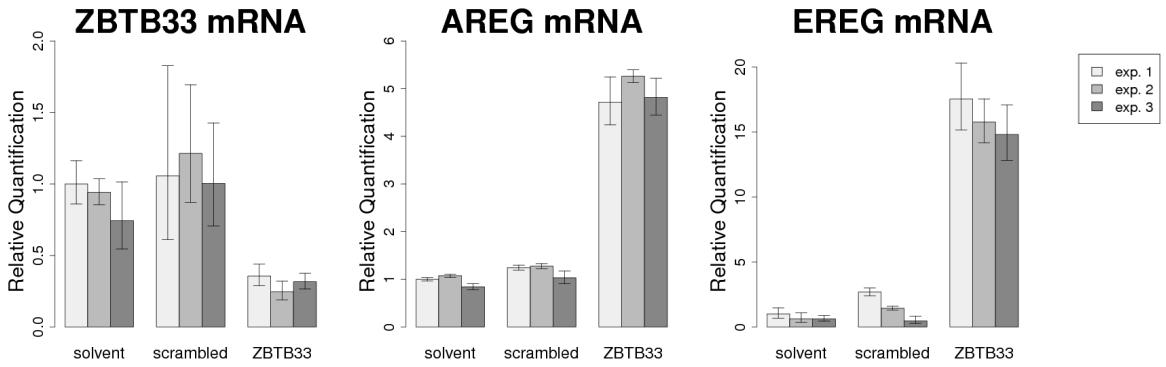


Figure 33: *AREG* and *EREG* expression in LIM1215 cells transfected with siRNAs targeting *ZBTB33*: LIM1215 cells were transfected transiently with siRNAs targeting *ZBTB33* in 3 independent experiments. As controls LIM1215 cells were either left untransfected (solvent) or were transfected with an unspecific siRNA (scrambled). Afterwards, RNA of these samples was converted into cDNA which was tested by real-time PCR for their *ZBTB33*, *AREG* and *EREG* expression.

3.8 An antisense transcript was addressed by Northern blot experiments and strand-specific PCR

In HCT116 cells, the *AREG* exon 2 sequence increased promoter function, when coupled in reverse orientation to the *AREG* promoter sequence or the Tata protein binding motif (see section 3.6). One possible explanation could be that *AREG* exon 2 stimulates the expression of a product, which is coded on the reverse complementary strand within the *AREG* gene. Such products could be RNAs like long non-coding RNAs (lncRNAs) or micro RNAs (miRNA) which might interfere with the *AREG* gene regulation. To test this hypothesis, a Northern blot experiment and a strand-specific PCR approach were performed.

3.8.1 Northern blot experiments

For the Northern blot experiment, RNA-probes were designed binding to both strands. The binding positions of the probes are shown in figure 34. The sequence of the RNA-probe "E2-sense" is reverse complementary to the *AREG* exon 2 sequence and therefore would bind to the *AREG* mRNA. This probe was used as a positive control. To address the potential promoter function of the reverse-oriented *AREG* exon 2, the "I2-antisense" probe was used for the experiment, because it would bind to a reverse-oriented RNA sequence starting at *AREG* exon 2. "E2-antisense" would also bind to a reverse-oriented RNA sequence. However, this sequence is within *AREG* exon 2, because it is not clear, where the sequence of the hypothesized ncRNA starts.

Figure 34:

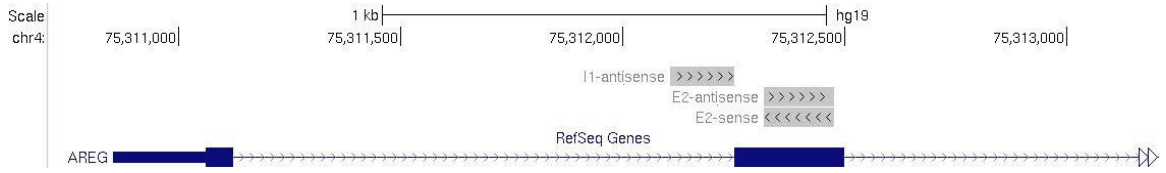


Figure 34: Binding sites of RNA-probes within the *AREG* gene: Three probes were designed to detect a hypothetical reverse complement non-coding RNA starting at exon 2. "E2-sense" binds within *AREG* exon 2 and should detect the *AREG* mRNA. "E2-antisense" would bind on the reverse complement sequence of *AREG* exon 2 and "I1-antisense" would bind on a reverse complement sequence in *AREG* intron 1 starting at exon 2.

In figure 35 the Northern blot results for poly-A RNA from 6 different cell lines including LIM1215, as well as LIM1215 treated with DAC are shown using the three probes described in figure 34. Three exposure times are shown to demonstrate that each probe created signals depending on the exposure time. The signals of the *AREG* mRNA detected by the "E2-sense" probe correlated well with the *AREG* mRNA expression detected by real-time PCR (see figure 3). Strong signals developed after short exposure in the lanes loaded with CaCO2-, HT29- and HCT116 poly-A RNA. A further signal

appeared after medium exposure in the lanes loaded with untreated and DAC-treated LIM1215 poly-A RNA. Very low signals appeared after long exposure in the lanes loaded with RKO- and SW480 poly-A RNA and also in the lane loaded with solvent treated LIM1215 poly-A RNA. The main difference in the Northern blot experiments using probe "E2-antisense" and "I1-antisense" was that it needed a long exposure of the "E2-antisense" blot to create a similar strong signal as appeared after short exposure of the "I1-antisense" blot. Since both probes were equally long and were used in the same amount, the difference might be due to different abundances of antisense binding sites. However, the binding sites of both probes to a hypothetical antisense RNA are in close proximity (see figure 34). The main similarity of both Northern blots was that the signals were spread over the whole blots and no individual band could be identified, representing the RNA. An argument could be that the signal spread was due to unspecific binding of the "E2-antisense" probe as well as the "I1-antisense" probe. However, the "E2-sense" blot shows that the length of the probes was sufficient to create a specific signal.

To sum up, the hypothetic antisense-transcript which might regulate *AREG* gene expression was not verified, because no specific signal appeared in the corresponding blots. However, a signal spread was observed. The validity of the Northern blot experiment is also reduced due to the discrepancy of the untreated LIM1215 RNA and the solvent treated LIM1215 RNA (lane 2 and 3). Although the effect points for a wrong amount of loaded RNA for this particular sample, a clear assumption could be made by introducing a loading control for all blots. Furthermore, the marker lanes also showed different signals by hybridizing with different probes which further reduces the validity of the experiment.

Figure 35:

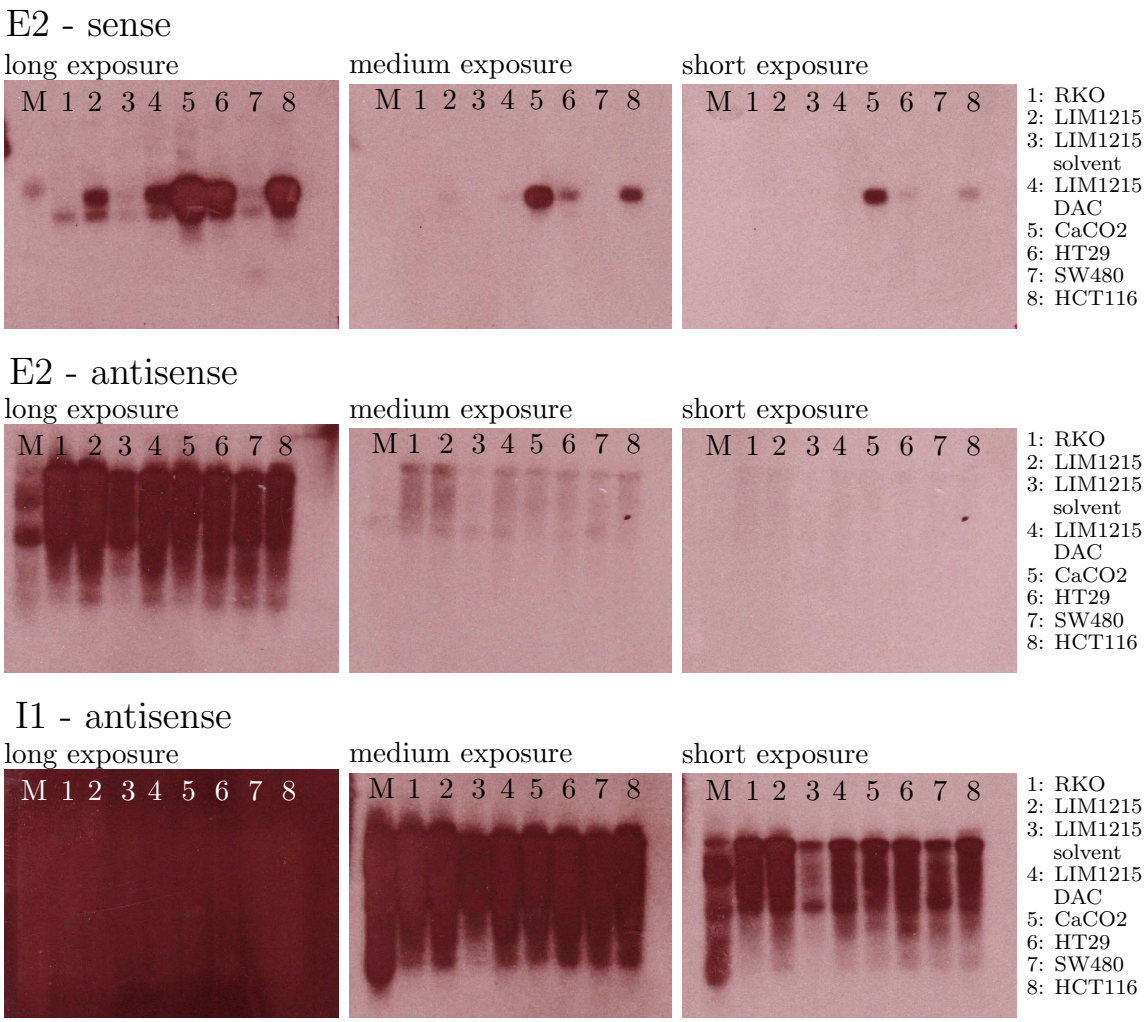


Figure 35: Northern blot analysis of several cell lines: Poly-A RNA of 6 different cell lines was loaded onto three gels. Northern blot was performed as described in section 2.10. The blots were incubated with the RNA-probes described in figure 34. Signals on a radiographic film are shown.

3.8.2 Strand-specific PCR

Because the Northern blot experiments only gave unclear results, a second approach was applied to clarify if there is an antisense RNA starting at *AREG* exon 2. At this PCR-based approach the cDNA obtained after reverse-transcription was converted using sodium bisulfite. All cytosines within the cDNA, which would normally bind to guanine in a double-stranded DNA, were converted into uracil, which would bind adenine. With this technique, it is possible to distinguish both strands of a double-stranded DNA product, or as used here to distinguish a cDNA derived from a sense-RNA as e.g. mRNA, from a cDNA derived from an antisense RNA of the same region. With PCR-primers specifically amplifying the antisense sequence, it is possible to detect the antisense RNA. An overview of this technique is shown in figure 36.

Figure 36:

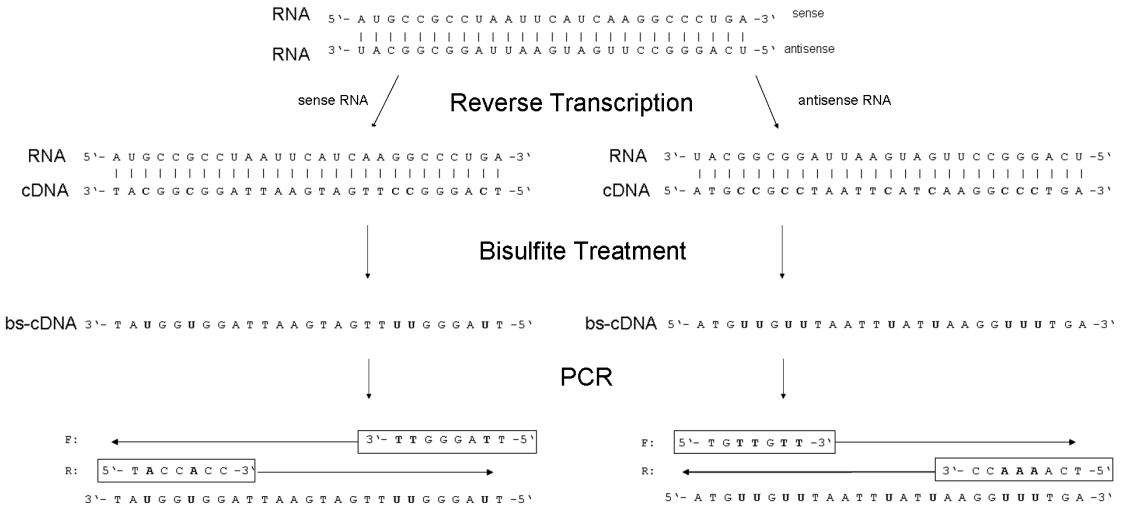
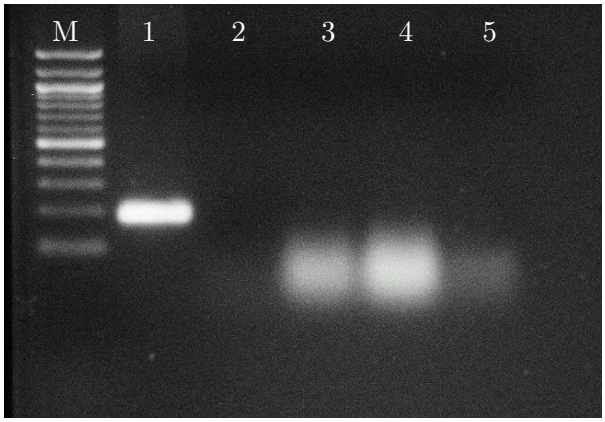


Figure 36: Scheme of strand-specific PCR: An RNA and its antisense RNA give different cDNAs, which could not be distinguished by standard PCR. Bisulfite conversion of the cDNAs leads to a change of all cytosines (C) to uracil (U). The cDNAs can now be distinguished by PCR using strand-specific primers.

An important requirement for using strand-specific PCR is that the RNA used for reverse transcription is free of any genomic DNA, since bisulfite-converted genomic DNA contains sequences of both strands and would lead to false-positive signals. A test PCR was performed, using primers amplifying a well-characterized genomic region within the *KRAS* gene. In figure 37 is shown that the genomic DNA (positive control) gave a clear signal (lane 1), which was not present in the lanes of the RNAs isolated from SW480, CaCO2 and LIM1215 (lanes 3-5) Only the absorbances of the RNAs itselves were visible in these lanes. The RNA was then used as template for reverse-transcription and the obtained cDNAs were bisulfite-converted using the Epitect-kit (Qiagen). To test, if the presence of the RNAs might inhibit PCR reaction using these primers, a DNA-contaminated RNA was used at a similar high concentration as template in another PCR. Here the same signal appeared as seen in the positive control (data not shown).

Figure 37:



1: genomic DNA
2: NTC
3: SW480 RNA
4: CaCO2 RNA
5: LIM1215 RNA

Figure 37: Contamination test of isolated RNA: RNA isolated from SW480 cells (3), CaCO2 cells (4) and LIM1215 cells (5) were used as template for a PCR amplifying a well-characterized locus within the *KRAS* gene. As a positive control HCT116-derived genomic DNA was used (1), A non-template control (NTC) was used as negative control (2).

Primers were designed to detect the bisulfite-converted cDNA derived from the *AREG* unspliced mRNA and the bisulfite-converted cDNAs derived from an antisense-RNA located within *AREG* exon 2 and *AREG* intron 1 (see figure 38). As controls, primers were designed to bind within two exons of the control gene *UBE2D2*. The chosen exons are separated by an 86 bp intron, which means that the amplification product confirmed the RNA, used for reverse-transcription, to be free of any residual genomic DNA. As a negative control primers were designed to amplify the same locus but on the reverse complementary strand. The results of the PCRs using bisulfite-converted cDNA derived from RNA isolated from LIM1215, SW480 and CaCO2 cells are shown in figure 39.

Figure 38:

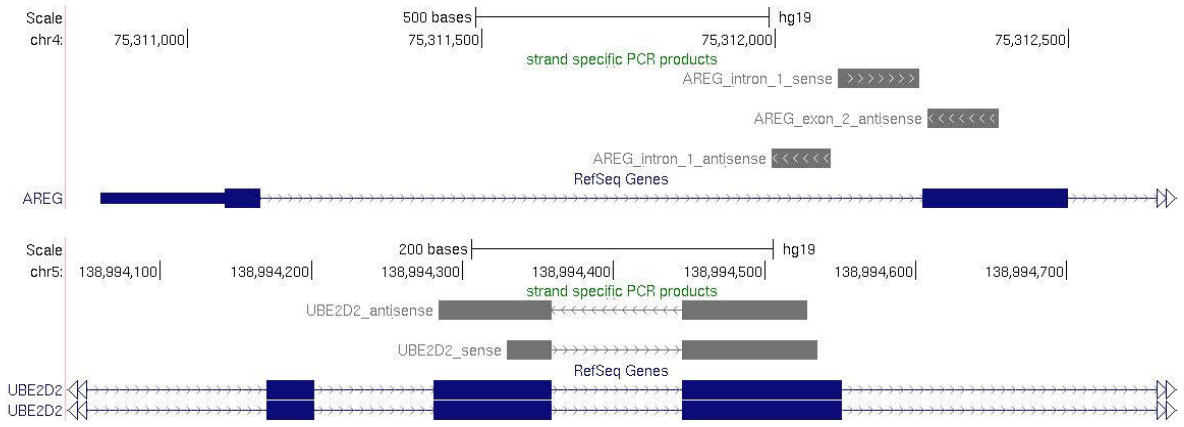


Figure 38: Location of the strand-specific PCR-products within the genome: Primerpairs were designed to detect the sense *AREG* intron 1 sequence as well as the antisense *AREG* exon 2 and intron 1 sequences. As controls, primerpairs were designed to detect sequences within the *UBE2D2* gene orientation dependently.

Figure 39:

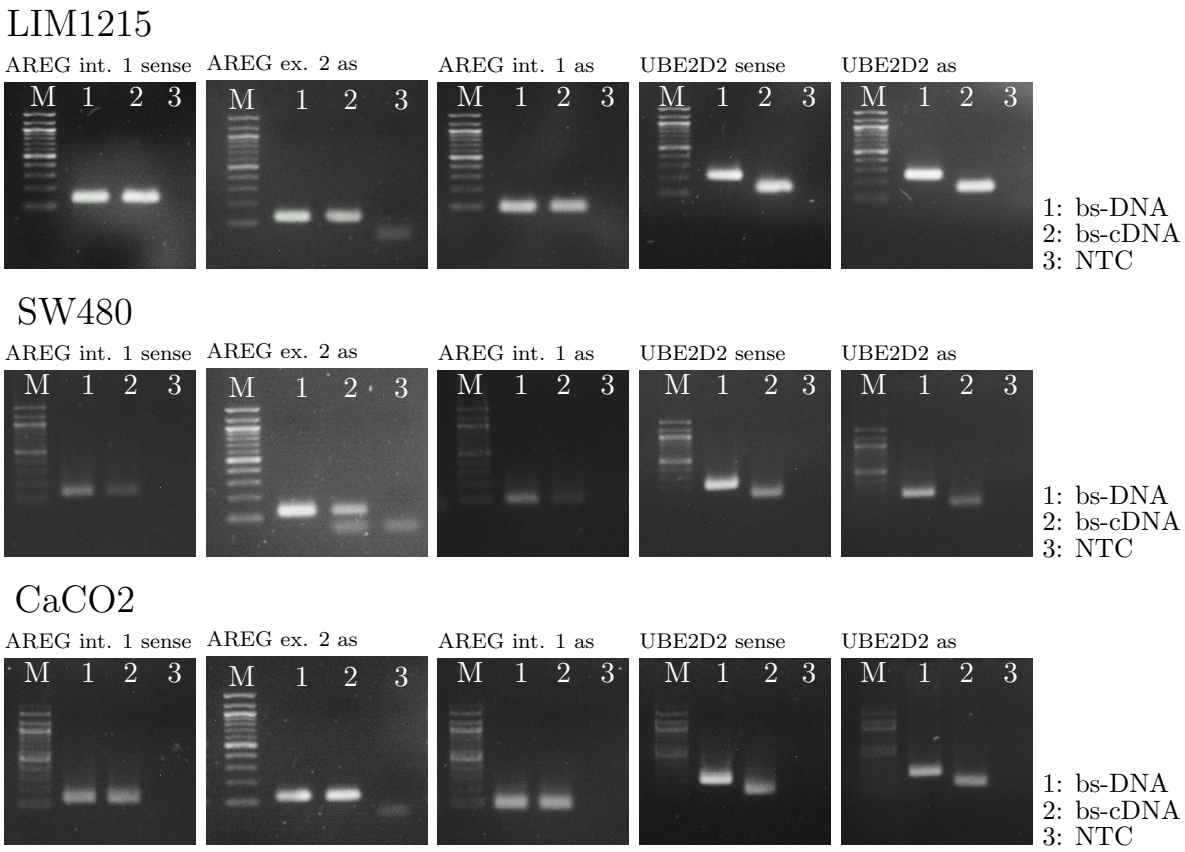


Figure 39: Strand-specific PCR: Bisulfite-converted cDNA (bs-cDNA) derived from LIM1215 cells, SW480 cells or CaCO2 cells was used as template to amplify the loci described in figure 38 Bisulfite-converted DNA, derived from HCT116^{DKO} cells (bsDNA) as well as a non-template control (NTC) were used as controls. M: 100 bp-marker.

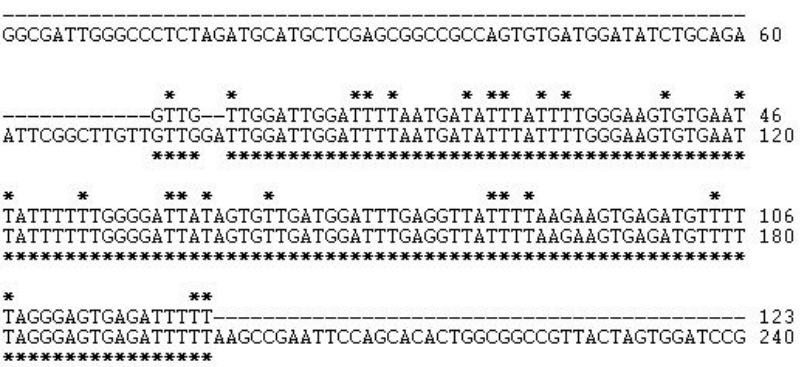
As shown in figure 39 the PCRs gave similar results for all three cell lines. It was confirmed by the signals showing the "UBE2D2 sense" amplification products that the template was free from genomic DNA, since the bisulfite-converted cDNA signal (bs-cDNA) appeared at approximately 150 bp, whereas the bisulfite-converted DNA signal (bs-DNA) appeared about 80 bp above the bs-cDNA signal. This was as expected. The amplification product detecting the unspliced *AREG* mRNA derived cDNA (AREG int. 1 sense) was present in all bs-cDNA samples and also as expected in all bs-DNA samples. The signal appeared as expected at approximately 150 bp. The PCRs, which should confirm an antisense transcript to the *AREG* gene (AREG int. 1 as and AREG ex. 2 as) showed signals in all three cell lines. These signals were similar to the signals appearing in the PCRs using bs-DNA as template and had also the expected size of approximately 100 bp and 150 bp, respectively. Therefore, the data strongly suggested an antisense transcript to the *AREG* gene. There was also a signal in the NTC of the PCR amplifying the AREG ex. 2 as. However, this signal laid below 100 bp and did not represent the amplification product. The signal might appear due to primer-dimerization, but did not influence the outcome. An unexpected signal was observed in the negative control PCR (UBE2D2 as). It should function as a negative control,

because its amplification product should only give signals in the presence of an antisense RNA to the control gene *UBE2D2*. However, it raised a signal in all three cell lines. To clarify, if this signal was unspecific, or if there is an antisense RNA within the *UBE2D2* gene and to show that the conclusions drawn for the *AREG* antisense RNA were still valid, the amplification products "UBE2D2 as", "AREG ex. 2 as" and "AREG int. 1 as" were sequenced using the TA cloning kit (Invitrogen). The amplification products came from bs-cDNA derived from the LIM1215 cell line A comparison of the sequencing results with the expected sequences is shown in figure 40.

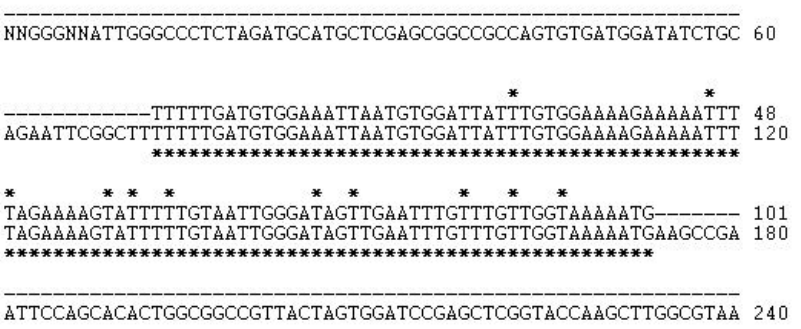
When sequencing the amplification products, it was observed that the sequences matched in all tested samples the expected sequences (see figure 40). The sequencing results also showed that the PCR was strand-specific as expected. All former cytosines of the sequence were converted to uracil by bisulfite-treatment which changed to thymine during PCR (marked by asterisks above the sequences in figure 40). Because the ligation reaction at the TA-cloning kit occurred orientation-independently, the thymine bases showed up at the sequence only, when the PCR product was ligated in forward orientation. This occurred for "AREG exon2 antisense" and "AREG intron 1 antisense". Since the samples were all free of contaminating genomic DNA, the results pointed for an *AREG* gene antisense transcript, which sequence contains a part of the reverse complement sequences of *AREG* intron 1 and *AREG* exon 2. The "UBE2D2 antisense"-PCR product was ligated by chance in reverse complement orientation into the TA-cloning vector. Therefore, the sequencing result represents the reverse complement sequence of the PCR-product. All thymine bases, derived from former cytosines are represented here as reverse complement adenines (also marked by asterisks in figure 40). Importantly, although the reverse complement sequence of the PCR product is shown here, the PCR product is not derived from the *UBE2D2* sense transcript. It rather verified an *UBE2D2* antisense transcript.

Figure 40:

AREG exon 2 antisense.



AREG intron 1 antisense



UBE2D2 antisense

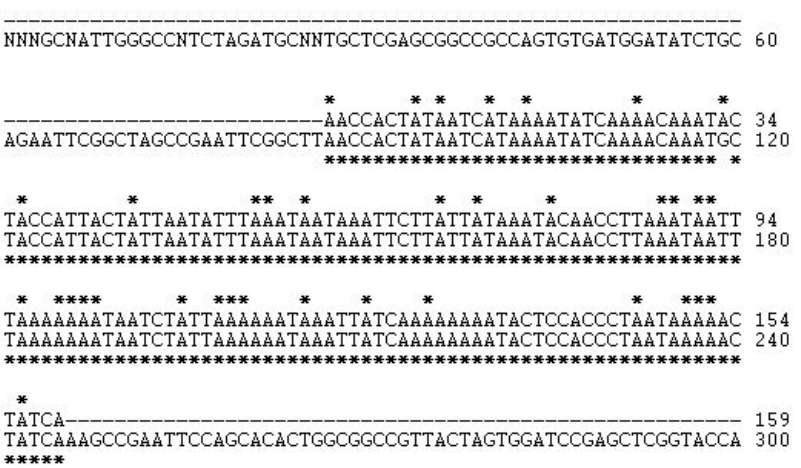


Figure 40: Sequencing results of strand-specific PCR products: bs-cDNA derived from LIM1215 was used as template for strand-specific PCR amplifying three loci. The PCR products were sequenced using the TA-cloning kit (Invitrogen). Sequencing results (upper lanes) were compared with expected sequences (lower lanes) using the webtool "ClustalW". Matched bases are indicated with asterisks below the sequences. Asterisks above the sequences indicate bisulfite-treatment dependent converted bases.

3.9 Epigenetic compounds can change the sensitivity of colorectal cancer cell lines towards EGFR inhibitors

In several studies, epigenetic compounds were tested for their therapeutic effects in cancer treatment. It was shown for lung cancer cells,[243] Hep-2 cancer cells (human larynx squamous cancer cell)[244] and SCCHN-cells,[245] that HDACis can lead to an increased sensitivity towards EGFR inhibitors like Erlotinib and Gefitinib. At the same time, observations by Khambata-Ford *et al.*[193] suggested a correlation between *AREG* and *EREG* expression and response of patients towards EGFR-targeted therapies. Therefore, it was tested next, if DAC or HDACis, while increasing *AREG* expression, may lead to an increase of sensitivity towards EGFR-inhibition in our cell lines, and afterwards, if the *AREG* expression is responsible for the increased sensitivity.

The sensitivity of different colorectal cancer cell lines upon Erlotinib and Gefitinib treatment was tested first.

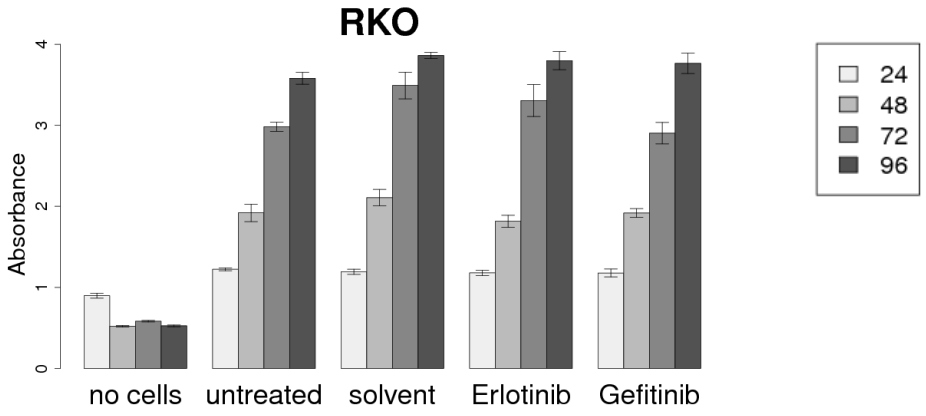
3.9.1 Sensitivity of untreated colorectal cancer cells towards EGFR-inhibition

Five cell lines, including two cell lines with low *AREG* expression (RKO, SW480), one cell line with medium *AREG* expression (LIM1215) and two cell lines with high *AREG* expression (HCT116, CaCO2), were tested for their response towards the EGFR inhibitors Erlotinib and Gefitinib. The phosphorylation of the EGFR and the phosphorylation of the downstream proteins AKT, MEK and ERK were tested by the Bio-Plex® technique (BioRad, see section 2.7) and the cell proliferation was tested by XTT proliferation assays. Besides the data shown here for Erlotinib and Gefitinib, the cells were also tested for their growth response towards Cetuximab treatment, a monoclonal antibody against EGFR and a common therapeutic compound in targeted therapies. However, all cells were fully resistant to Cetuximab *in vitro* which was in contrast to observations made by Jhawer *et al.*,[246] (data not shown). Therefore, Cetuximab was omitted from further analysis.

Figure 41:

RKO

A: Growth Determination



B: Protein Phosphorylation

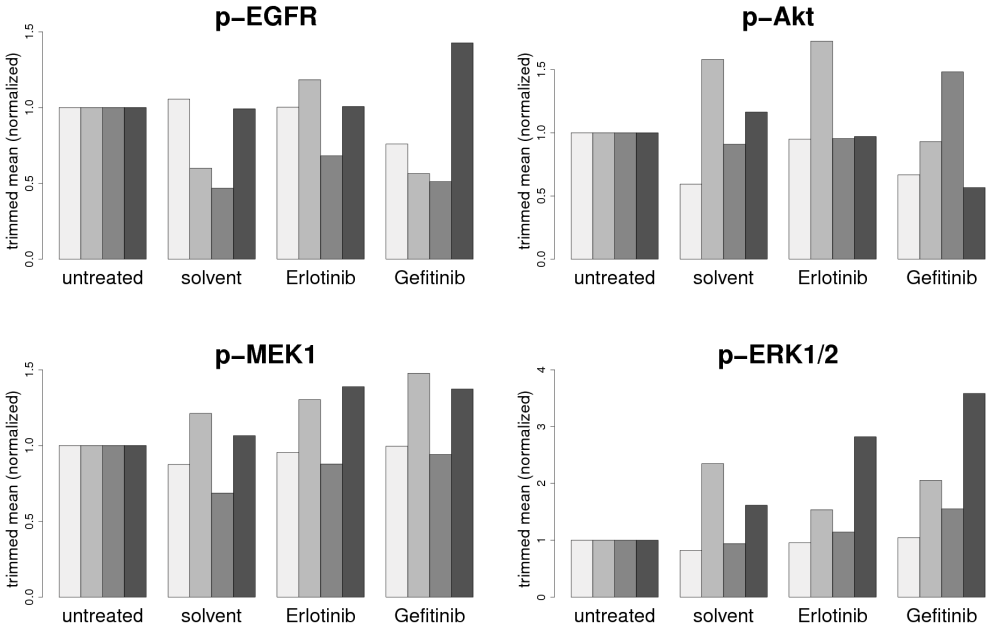
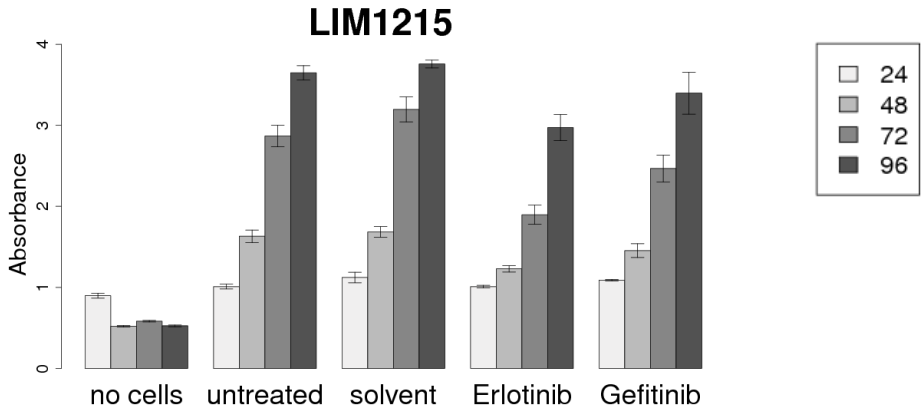


Figure 41: Characterization of RKO cells after EGFR inhibitor treatment: RKO cells were treated with 10 µM Erlotinib or 10 µM Gefitinib for 3 or 4 timepoints A: Growth was measured using the XTT-assay: Treatment was performed daily starting 24 h after seeding and continuing the next 96 h (see 2.3.5). Absorbance at 480 nm was measured 24 h after XTT solution dispense. Mean values plus standard deviations of three replicates are shown. B: Phosphorylation of EGFR and downstream proteins were tested by the Bio-Plex® technique (see section 2.7).

Figure 42:

LIM1215

A: Growth Determination



B: Protein Phosphorylation

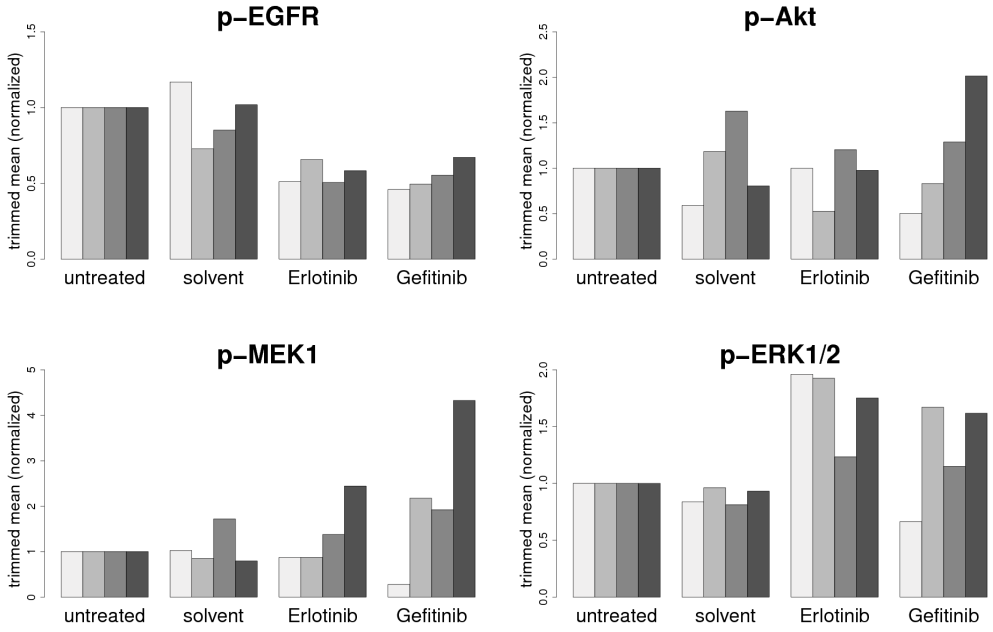
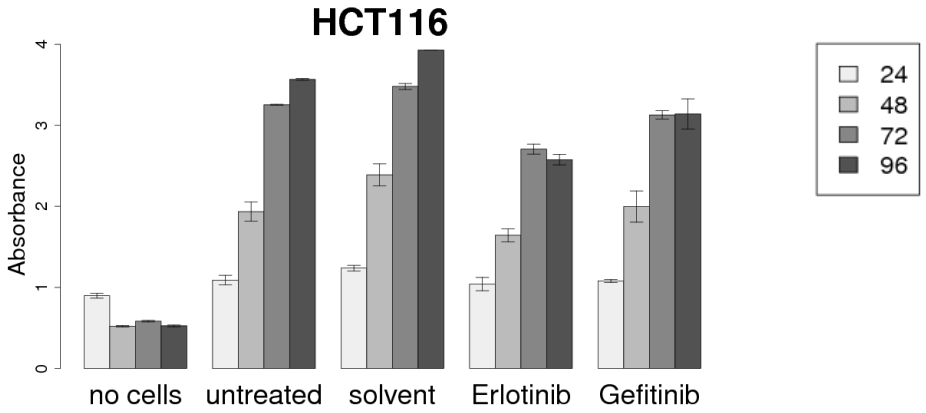


Figure 42: Characterization of LIM1215 cells after EGFR inhibitor treatment: LIM1215 cells were treated with 10 µM Erlotinib or 10 µM Gefitinib for 3 or 4 timepoints. A: Growth was measured using the XTT-assay: Treatment was performed daily starting 24 h after seeding and continuing the next 96 h (see 2.3.5). Absorbance at 480 nm was measured 24 h after XTT solution dispense. Mean values plus standard deviations of three replicates are shown. B: Phosphorylation of EGFR and downstream proteins were tested by the Bio-Plex® technique (see section 2.7).

Figure 43:

HCT116

A: Growth Determination



B: Protein Phosphorylation

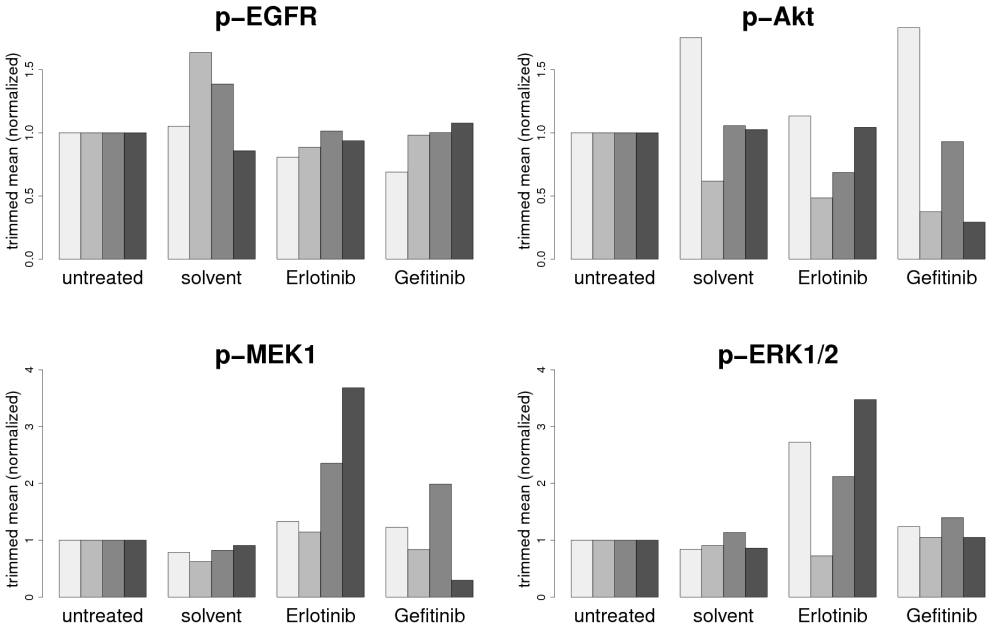
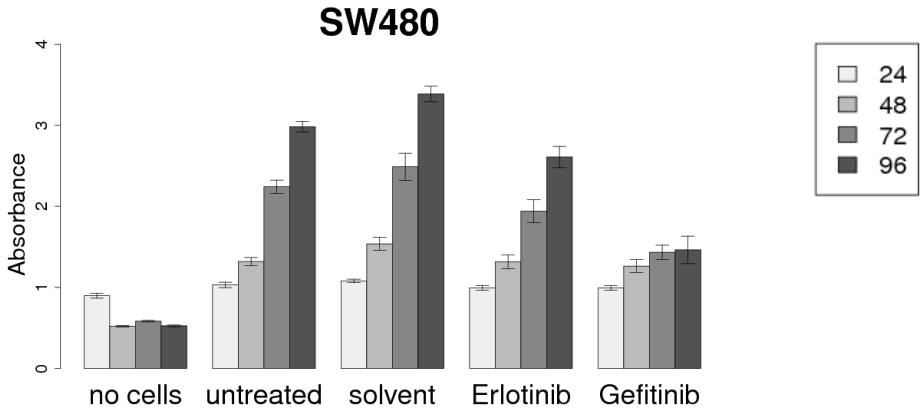


Figure 43: Characterization of HCT116 cells after EGFR inhibitor treatment: HCT116 cells were treated with 10 µM Erlotinib or 10 µM Gefitinib for 3 or 4 timepoints. A: Growth was measured using the XTT-assay: Treatment was performed daily starting 24 h after seeding and continuing the next 96 h (see 2.3.5). Absorbance at 480 nm was measured 24 h after XTT solution dispense. Mean values plus standard deviations of three replicates are shown. B: Phosphorylation of EGFR and downstream proteins were tested by the Bio-Plex® technique (see section 2.7).

Figure 44:

SW480

A: Growth Determination



B: Protein Phosphorylation

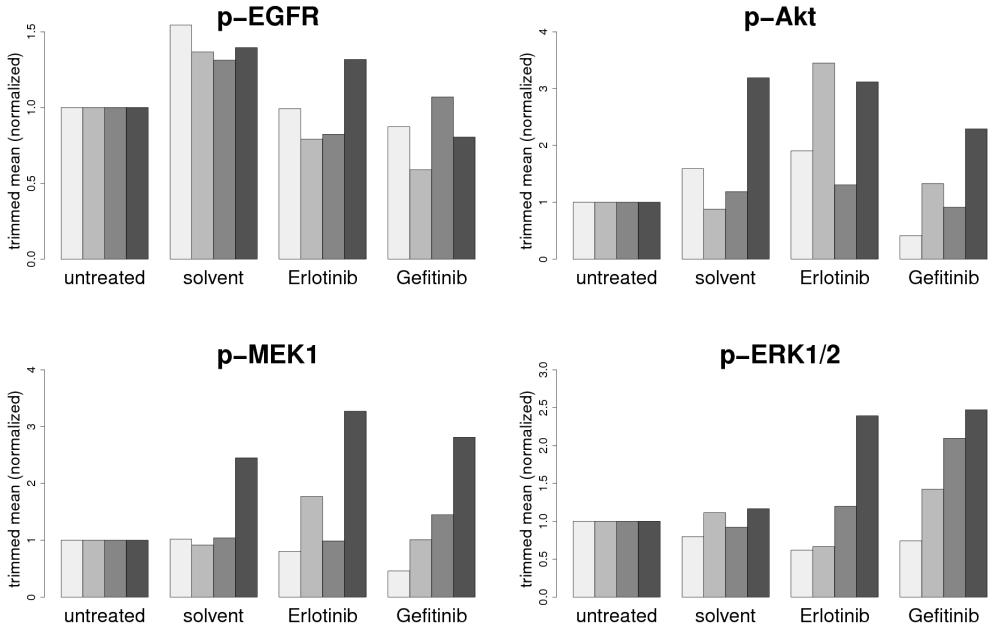
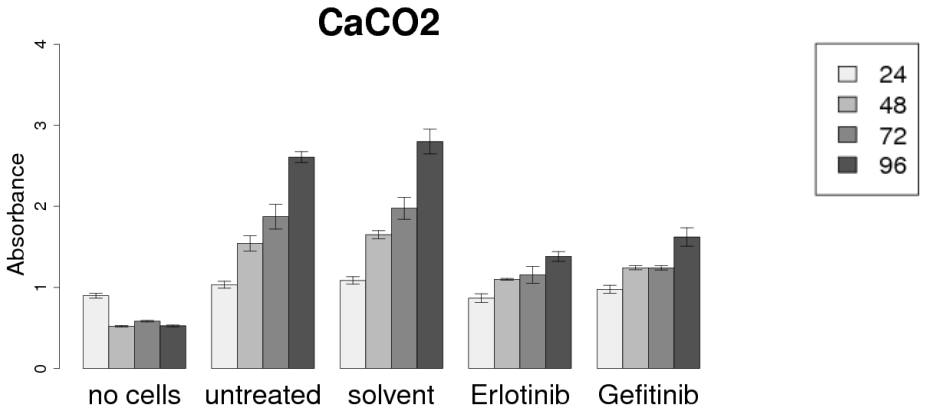


Figure 44: Characterization of SW480 cells after EGFR inhibitor treatment: SW480 cells were treated with 10 µM Erlotinib or 10 µM Gefitinib for 3 or 4 timepoints. A: Growth was measured using the XTT-assay: Treatment was performed daily starting 24 h after seeding and continuing the next 96 h (see 2.3.5). Absorbance at 480 nm was measured 24 h after XTT solution dispense. Mean values plus standard deviations of three replicates are shown. B: Phosphorylation of EGFR and downstream proteins were tested by the Bio-Plex® technique (see section 2.7).

Figure 45:

CaCO2

A: Growth Determination



B: Protein Phosphorylation

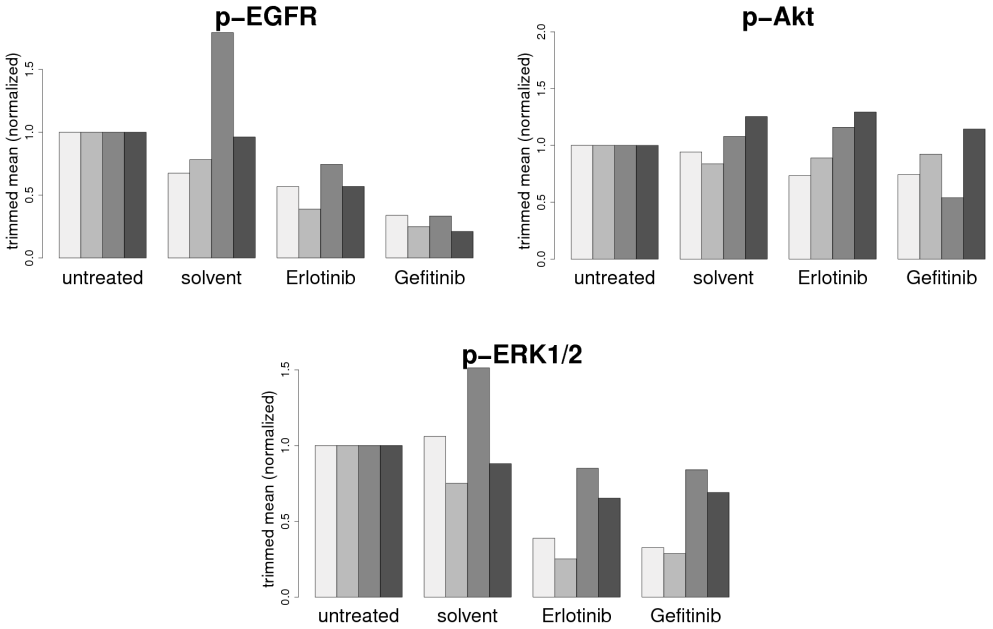


Figure 45: Characterization of CaCO2 cells after EGFR inhibitor treatment: CaCO2 cells were treated with 10 µM Erlotinib or 10 µM Gefitinib for 3 or 4 timepoints. A: Growth was measured using the XTT-assay: Treatment was performed daily starting 24 h after seeding and continuing the next 96 h (see 2.3.5). Absorbance at 480 nm was measured 24 h after XTT solution dispense. Mean values plus standard deviations of three replicates are shown. B: Phosphorylation of EGFR and downstream proteins were tested by the Bio-Plex® technique (see section 2.7).

When treating the RKO cell line with 10 µM Erlotinib or Gefitinib, no reduction in growth was observed (see figure 41 A). Therefore RKO is resistant to EGFR-inhibition. Interestingly, EGFR-phosphorylation increased for Erlotinib treated cells compared to the solvent treated cells after 48 h of treatment and for Gefitinib treated cells after 96 h of treatment (see figure 41 B). The EGFR-dephosphorylation, which should follow EGFR-inhibition, was compensated by an unknown mechanism. The mechanism might involve the MAPK-pathway, since ERK1/2-phosphorylation increased in the Erlotinib and Gefitinib treated cells. The phosphorylation states of the other proteins fluctuated at the different timepoints within the experiment.

When treating LIM1215 cells as well as HCT116 cells with Erlotinib or Gefitinib, a small growth reduction was observed for both cell lines (see figures 42 A and 43 A). In LIM1215 cells, growth reached for Erlotinib approximately 60 % and for Gefitinib approximately 80 % of the value of the solvent treated cells at 72 h. Nevertheless, growth continued to 96 h. In a similar way, HCT116 cells had a reduced growth at 72 h (reduction of approximately 20 % for Erlotinib, 10 % for Gefitinib). Different to LIM1215 cells, Erlotinib and Gefitinib treated HCT116 cells stopped growing after 72 h. In both cell lines a reduced EGFR-phosphorylation was seen upon Erlotinib and Gefitinib treatment, compared to the solvent control (see figures 42 B and 43 B). Interestingly, although EGFR was dephosphorylated upon treatment, an increased phosphorylation in MEK1 and ERK1/2 was observed in LIM1215 cells at the late timepoints for both Erlotinib and Gefitinib and in AKT for Gefitinib treated cells. In HCT116 cells, Erlotinib treatment also led to an increased phosphorylation of MEK1 and ERK1/2.
A strong growth reduction upon Gefitinib treatment was observed in SW480 cells (see figures 44 A). But the cells remained growing upon Erlotinib treatment. EGFR-phosphorylation decreased slightly for both Erlotinib and Gefitinib treated cells compared to the solvent treated cells (see figures 44 B). While phoshorylation of AKT increased after 24 and 48 h of Erlotinib treatment, a decrease of phosphorylation was observed in Gefitinib treated cells after 24 h and a similar phosphorylation was observed after 48 h. In contrast ERK1/2 phosphorylation was stronger affected after Gefitinib treatment at the early timepoints. However, at the late timepoints, phosphorylation of MEK1 (96h) and ERK1/2 (72-96h) is higher than in the solvent treated cells. Phosphorylation of AKT after Erlotinib and Gefitinib treatment matched the phosphorylation of AKT of the solvent treated cells at this timepoints.
By comparing solvent treated and Erlotinib/Gefitinib-treated CaCO2 cells, a strong growth reduction was observed for both inhibitors (see figure 45 A). Also similar to the other cells a reduction in EGFR-phosphorylation was observed (see figure 45 B). In contrast to the other cells, phosphorylation of ERK1/2 decreased for all timepoints. The phosphorylation of AKT remained almost constant in the whole experiment and was not influenced by Erlotinib or Gefitinib treatment. To summarize, the five tested cell lines behaved different upon Erlotinib and Gefitinib treatment. Although a dephos-

phorylation of the EGFR was mainly seen in all cell lines, downstream proteins showed various phosphorylation patterns. An increase of phosphorylation of the MAPK proteins MEK1 and ERK1/2 was seen in 4 of the 5 tested cells. Only CaCO2, which was also the only cell line having no mutations within the MAPK and PI3K-pathways, showed a reduction in ERK1/2 phosphorylation. Interestingly, this cell line was also the most sensitive cell line towards Erlotinib and Gefitinib treatment.

3.9.2 Sensitivity of cell lines after treatment with DAC and HDACi

To address the influence of epigenetically interfering compounds, experiments were designed as follows: The cells were first treated with the epigenetically interfering compound for 144 h (see section 2.3.4). Afterwards, the cells were collected and replated in equal numbers into 96-well plates. Then they were either left untreated or treated with Erlotinib, Gefitinib, or DMSO as a solvent control (see section 2.3.5). This treatment will be referred to as "post-treatment" on the following pages. Cell growth was evaluated by XTT experiments.

Figure 46:

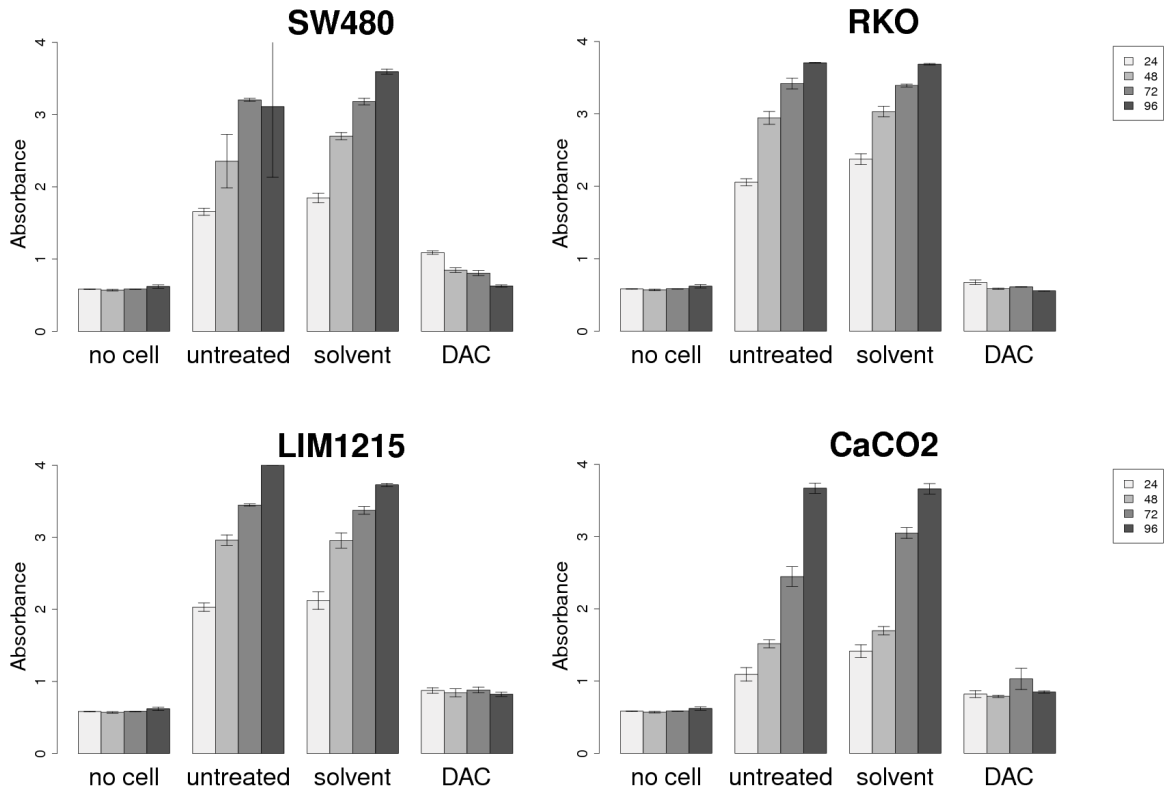


Figure 46: Influence of DAC treatment on cell growth: Four cell lines were treated with 2.5 µM DAC for 144 h as described in section 2.3.5. After transfer of equal cell numbers into a 96-well plate, cell growth was evaluated by XTT experiments. XTT absorbance was measured 24 h after XTT solution dispense.

To test, whether DAC has an effect onto the cells' sensitivity towards EGFR inhibitors, four cell lines were treated for 144 h with 2.5 µM DAC. After replating into 96-well plates, cell proliferation was monitored. As described in figure 46, none of the DAC-treated cells started growing after replating. Therefore, the effect of DAC onto sensitivity towards EGFR inhibitors could not be examined in that way. Lower concentrations of DAC were used for treatment, too (down to 0.5 µM). But also in these experiments, no growth was observed after DAC-treatment (data not shown).

Figure 47:

LIM1215:

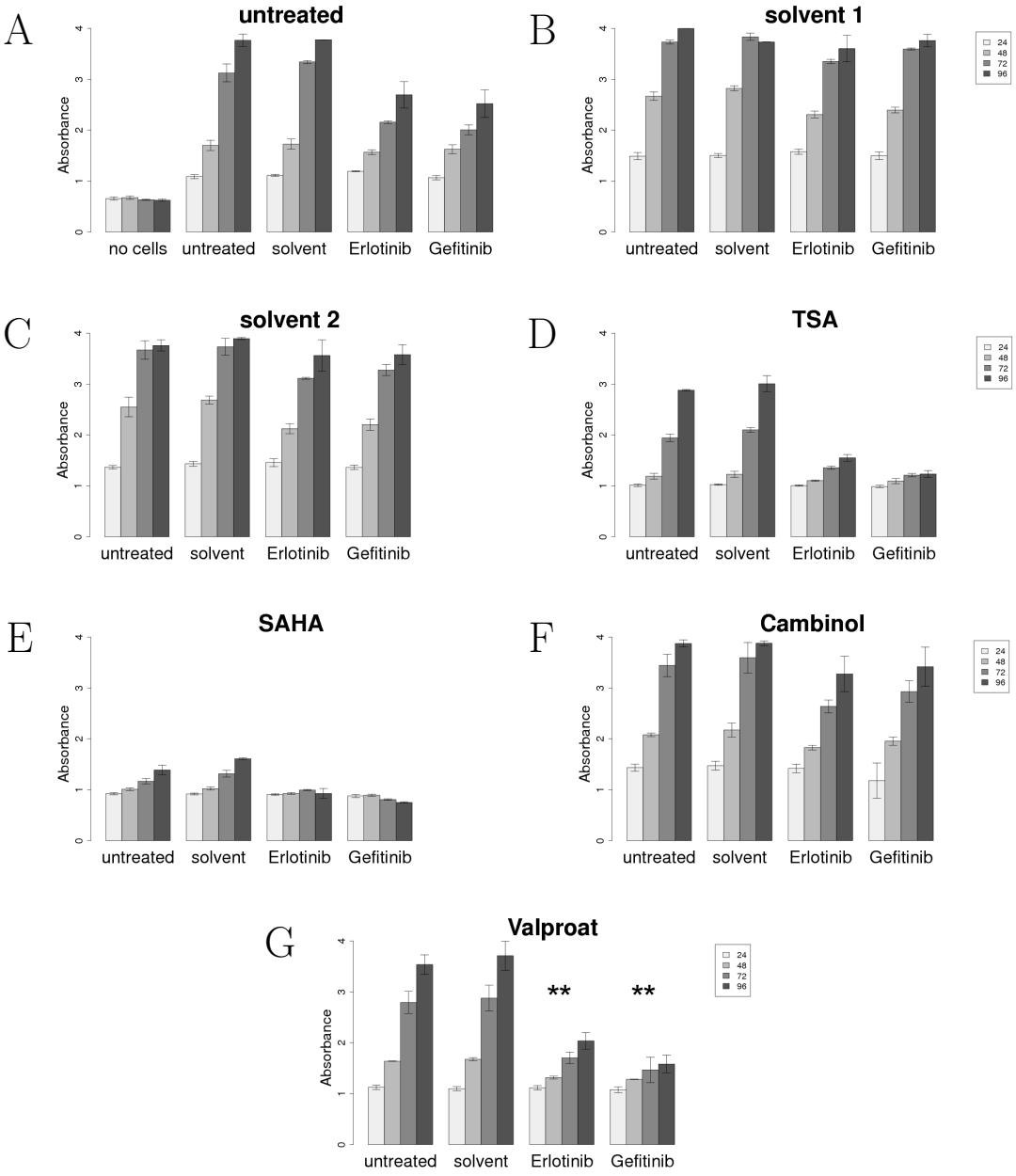


Figure 47: XTT proliferation experiments in LIM1215 cells after HDACi treatment: LIM1215 cells were treated with 4 different HDACis and their solvents for 144 h (TSA: 25 ng/ml, SAHA: 1 µM, Cambinol: 20 µM, Valproat: 1 mM, solvent 1: DMSO, solvent 2: Methanol). The treated cells were then replated into 96-well plates for XTT experiments. After 24 h, cells were treated daily with Erlotinib, Gefitinib and the solvent for an additional 4 days. Each day, one plate was used for XTT measurement. XTT absorbance was measured 24 h after XTT solution dispense. Statistical test: A heteroscedastic t-test was performed as described in the text *: the p-value is less than 0.05 at 96 h, **: the p-value is less than 0.01 upon condition that there is no significant difference at 96 h between the untreated or solvent post-treated samples after treatment with the HDACi and the solvent 1 or 2.

Figure 48:

SW480:

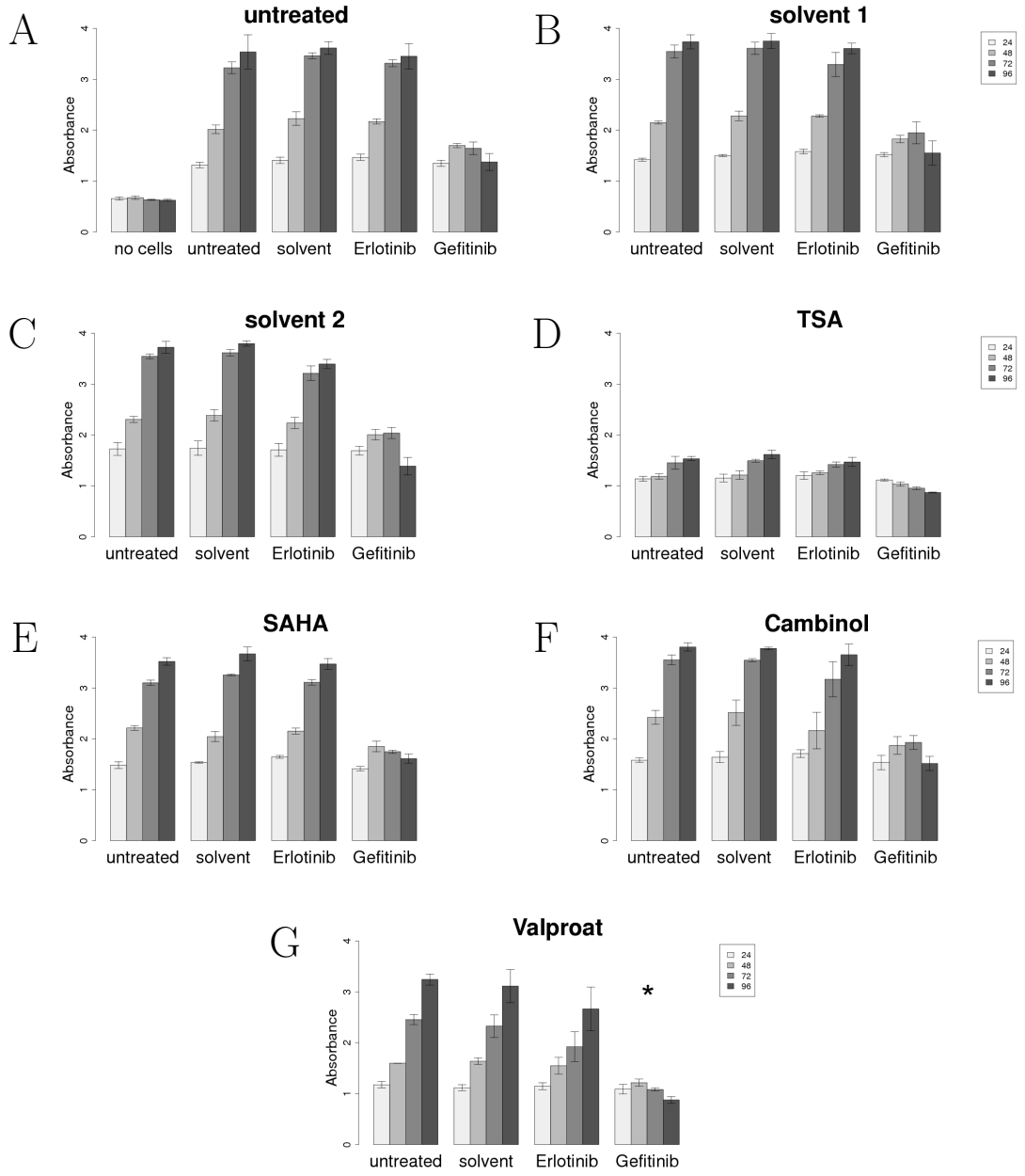


Figure 48: XTT proliferation experiments in SW480 cells after HDACi treatment: SW480 cells were treated with 4 different HDACis and their solvents for 144 h (TSA: 25 ng/ml, SAHA: 1 µM, Cambinol: 20 µM, Valproat: 1 mM, solvent 1: DMSO, solvent 2: Methanol). The treated cells were then replated into 96-well plates for XTT experiments. After 24 h, cells were treated daily with Erlotinib, Gefitinib and the solvent for an additional 4 days. Each day, one plate was used for XTT measurement. XTT absorbance was measured 24 h after XTT solution dispense. Statistical test: A heteroscedastic t-test was performed as described in the text *: the p-value is less than 0.05 at 96 h, **: the p-value is less than 0.01 upon condition that there is no significant difference at 96 h between the untreated or solvent post-treated samples after treatment with the HDACi and the solvent 1 or 2.

Similar to DAC-treated cells, HDACi-treated cells were tested for their growth after replating and if the HDACis have an influence onto sensitivity towards Erlotinib and Gefitinib. The response of LIM1215 and SW480 with or without HDACi treatment is shown in figure 47 and figure 48. Statistical tests were made as follows to address the effects of the HDACis onto the sensitivity. A heteroscedastic t-test was performed to compare (1.) the absorbances of the Erlotinib or Gefitinib post-treated samples after HDACi-treatment with the absorbances of the Erlotinib or Gefitinib post-treated sample after solvent treatment. The t-test was performed for the timepoint 96 h, because here the strongest effect should be seen. Additionally a heteroscedastic t-test was performed to compare (2.) the absorbances of the post-treatment solvent controls (solvent/untreated) after HDACi-treatment with the absorbances of the post-treatment solvent controls after solvent treatment. A significant difference in the first t-test indicates that sensitivity towards Erlotinib or Gefitinib depends on the HDACi treatment. However, a significant difference in the second t-test indicates that the cell growth is also affected by the HDACi itself. Only when the first test shows significant reduction and the second does not, asterisks are drawn to emphasize the significant effect of the HDACi on the sensitivity which is independent to growth related effects.

In contrast to DAC-treatment, both cell lines restarted proliferation after replating, when treated with HDACis. One exception in LIM1215 cells was SAHA-treatment (see figure 47 E) and another exception in SW480 cells was TSA-treatment (see figure 48 D). These exceptions were omitted from analysis. Interestingly solvent treated LIM1215 cells (DMSO and Methanol) were more resistant towards Erlotinib and Gefitinib than untreated cells (see figure 47 B-C). However, compared to these solvent controls, treatment with TSA and Valproat led to a higher sensitivity towards Erlotinib and Gefitinib in LIM1215 cells, which was highly significant for Valproat-treatment. Reduced growth curves occurred compared to the post-treatment solvent and untreated controls (see figure 47 D and G). Since the TSA-treated controls, solvent and untreated, showed also reduced growth curves to the solvent 1 treated controls (compare figures 47 B and D), it is obvious that TSA also had an effect on the growth of LIM1215 cells, and that the effects, seen after Erlotinib and Gefitinib might additionally be influenced by this TSA-effect. The HDACi Cambinol did not change sensitivity of LIM1215 cells towards EGFR-inhibition (see figure 47 F).
In Valproat-treated SW480 cells, a significant decrease of growth was observed after Gefitinib post-treatment (see figure 48 G). Therefore, similar to LIM1215 cells, Valproat led to an increased sensitivity towards Gefitinib in SW480 cells. However, it is unclear, why sensitivity towards Erlotinib is not influenced by Valproat. Besides this exception, all treated and replated cells were as resistant to Erlotinib and as sensitive to Gefitinib as the untreated control (compare figure 48 A with figures 48 B, C, E, F).

In both cell lines, Valproat increased response towards EGFR inhibitors (see figures 47 G and 48 G). Therefore, the following experiments were performed with Valproat.

Figure 49:

SW480:

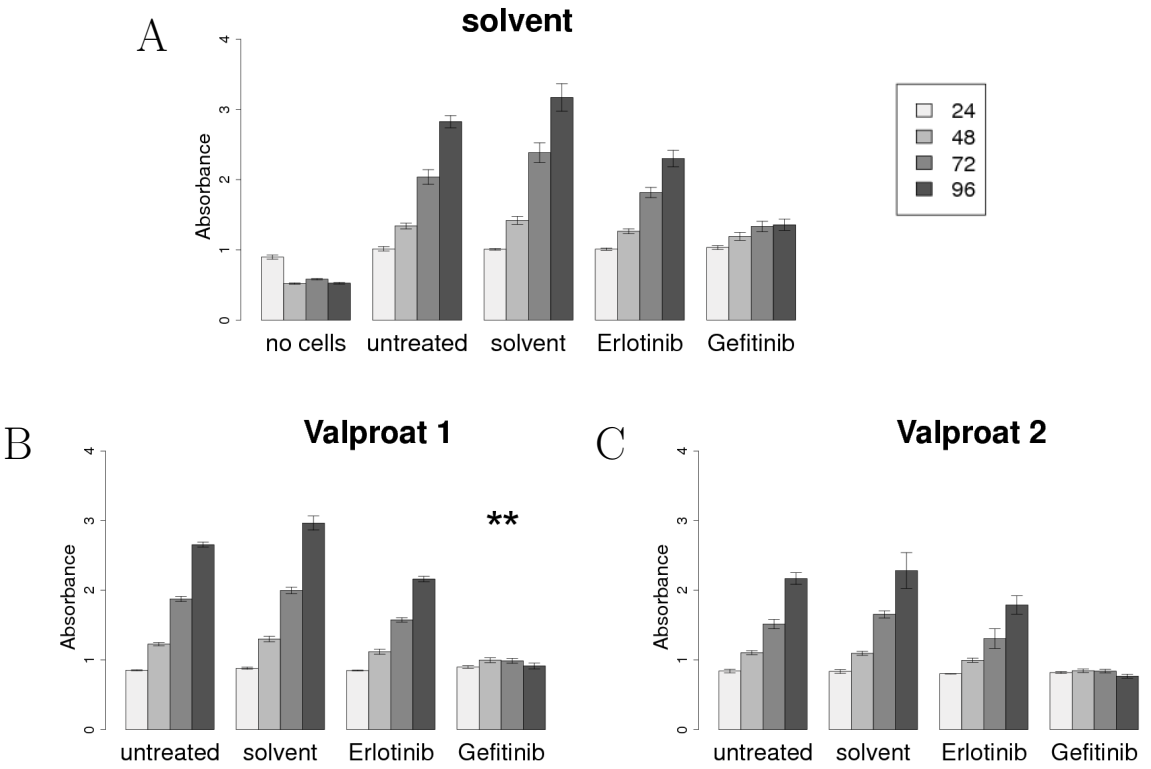


Figure 49: XTT proliferation experiments in SW480 cells after Valproat treatment: SW480 was treated with 1 mM or 2 mM Valproat (Valproat 1 or 2) or the solvent Methanol for 144 h. The treated cells were replated into 96-well plates for XTT experiments as described earlier. XTT absorbance was measured 24 h after XTT solution dispense. Statistical test: A heteroscedastic t-test was performed as described in the text *: the p-value is less than 0.05 at 96 h, **: the p-value is less than 0.01 upon condition that there is no significant difference at 96 h between the untreated or solvent post-treated samples after treatment with Valproat 1 or 2 and Methanol.

To test, if Valproat might have an effect on other cell lines and to test if a higher concentration of Valproat might trigger an effect, SW480, RKO, LIM1215, HCT116 and CaCO2 cells were treated for 144 h with 1 mM and 2 mM of Valproat. Afterwards, these cells were tested for their sensitivity towards Erlotinib and Gefitinib by XTT proliferation experiments as described earlier.

Similar to the experiment before, Valproat treatment significantly increased sensitivity towards Gefitinib but not towards Erlotinib in SW480 cells (see figure 49 B). Using a higher concentration of Valproat, an overall reduction of XTT absorbances was seen (figure 49 C). Therefore, using this concentration a Valproat-derived reduction in growth could not be distinguished from a EGFR inhibitor-derived reduction in growth.

Figure 50:

RKO:

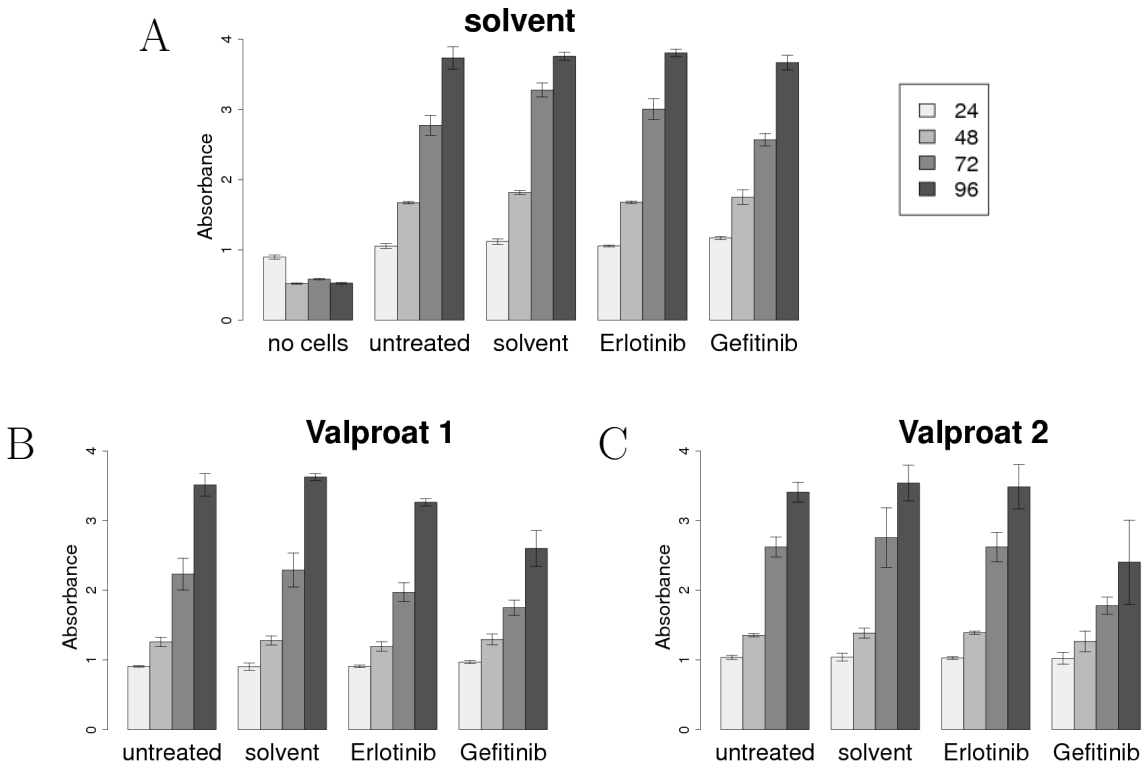


Figure 50: XTT experiments in RKO after Valproat treatment: RKO cells were treated with 1 mM or 2 mM Valproat (Valproat 1 or 2) for 144 h. The treated cells were replated into 96-well plates for XTT experiments as described earlier. XTT absorbance was measured 24 h after XTT solution dispense. Statistical test: The statistical test was performed as described in figure 49.

In figure 50 the effect of Valproat treatment onto EGFR-inhibition is shown for the RKO cell line. A minimal reduction in growth was observed after post-treatment with Gefitinib in Valproat treated cells irrespective of the Valproat concentration (see figure 50 B-C) However, this growth-reduction was not significant.

Figure 51:

LIM1215:

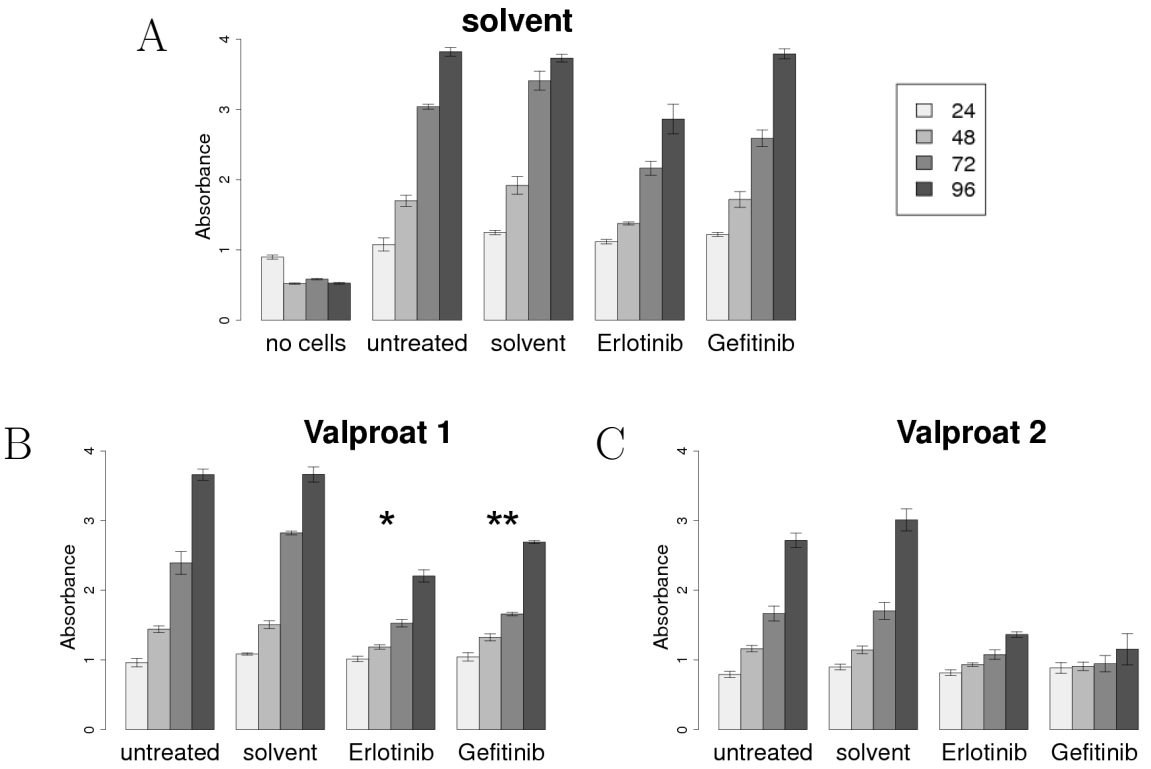


Figure 51: XTT experiments in LIM1215 after Valproat treatment: LIM1215 cells were treated with 1 mM or 2 mM Valproat (Valproat 1 or 2) for 144 h. The treated cells were replated into 96-well plates for XTT experiments as described earlier. XTT absorbance was measured 24 h after XTT solution dispense. Statistical test: The statistical test was performed as described in figure 49.

It was already shown that LIM1215 cells had a significantly increased sensitivity towards Erlotinib and Gefitinib upon Valproat treatment (see figure 47). After repeating the experiment, the increase of sensitivity was observed again (compare figure 51 A and B). Using 2 mM Valproat, a similar overall reduction in growth was observed in the LIM1215 cell line as seen in the SW480 cell line (see figure 51 C). Although the cells grew only very slow upon Erlotinib post-treatment or almost stopped growing upon Gefitinib post-treatment, the effect might be influenced by Valproat itself, since growth reduction was also seen for the post-treatment solvent controls.

Figure 52:

HCT116:

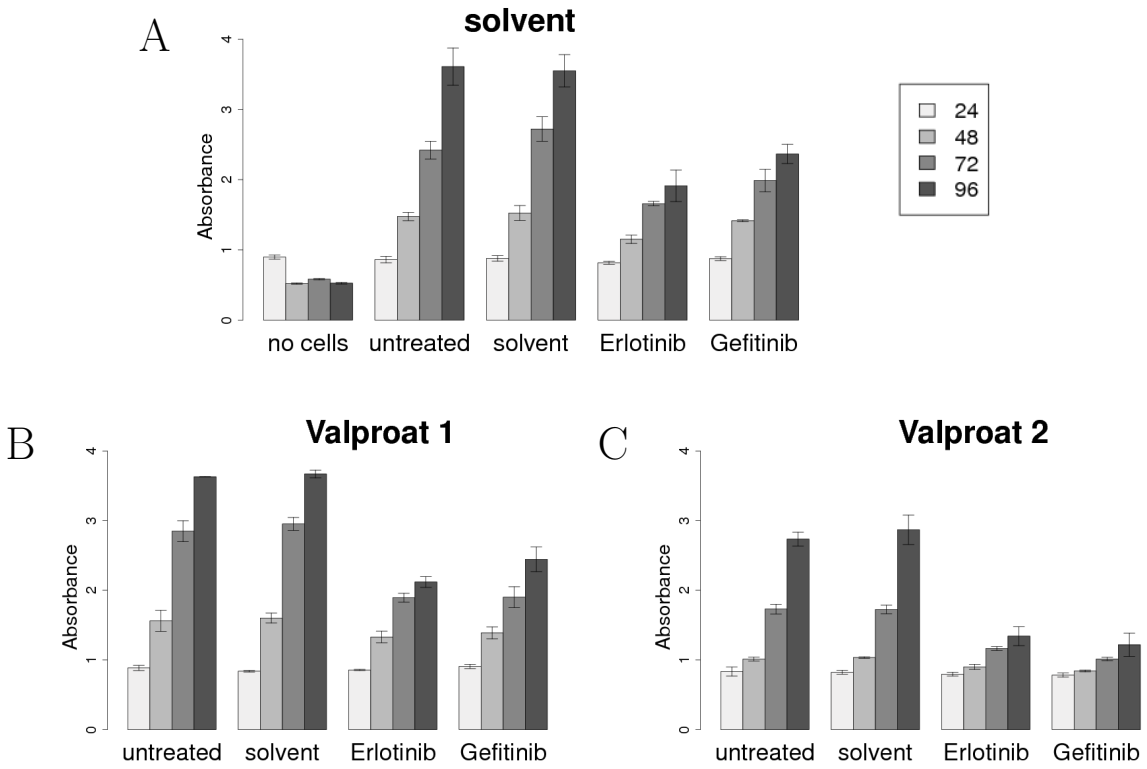


Figure 52: XTT experiments in HCT116 after Valproat treatment: HCT116 cells were treated with 1 mM or 2 mM Valproat (Valproat 1 or 2) for 144 h. The treated cells were replated into 96-well plates for XTT experiments as described earlier. XTT absorbance was measured 24 h after XTT solution dispense. Statistical test: The statistical test was performed as described in figure 49.

The results of HCT116 cells treated with Valproat are shown in figure 52. The solvent treated control cells were sensitive towards the EGFR inhibitors (see figure 52 A). When treating with 1 mM Valproat, the XTT aborbances of the post-treatment experiments were similar to the solvent treated cells (see figures 52 A and B). The sensitivity towards Erlotinib and Gefitinib was not increased. But, when treating the cells with 2 mM Valproat, the Erlotinib and Gefitinib post-treated cells grew slower (see figure 52 C). However, the post-treatment controls (untreated and solvent treated) had also a reduced growth rate compared to the solvent treated HCT116 cells (figure 52 A). Therefore, an effect of Valproat itself onto growth of HCT116 cells could not be ruled out.

Figure 53:

CaCO2:

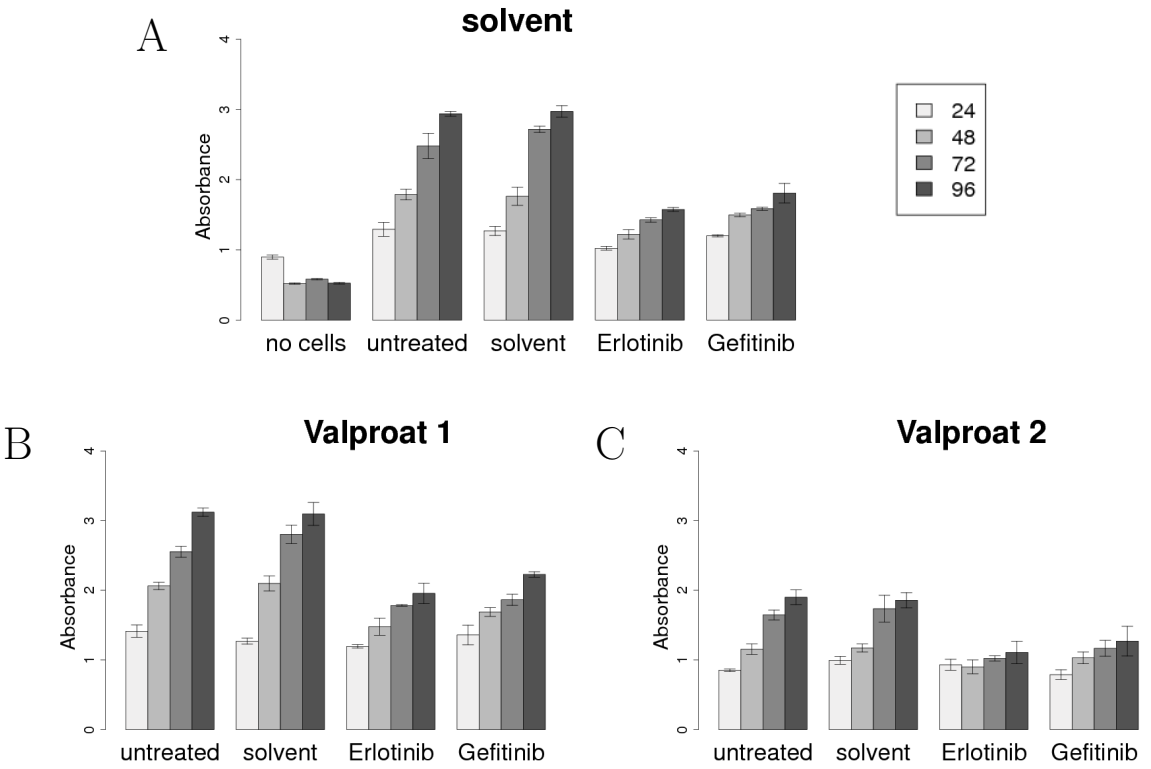


Figure 53: XTT experiments in CaCO2 after Valproat treatment: CaCO2 cells were treated with 1 mM or 2 mM Valproat (Valproat 1 or 2) for 144 h. The treated cells were replated into 96-well plates for XTT experiments as described earlier. XTT absorbance was measured 24 h after XTT solution dispense. Statistical test: The statistical test was performed as described in figure 49.

The sensitivity of CaCO2 cells was not increased upon 1 mM or 2 mM Valproat treatment. The aborbances were similar to the solvent treated control cells. After 2 mM Valproat treatment, CaCO2 grew much slower even in the absence of Erlotinib or Gefitinib. As a consequence, similar to HCT116 cells, a Valproat-mediated effect could not be distinguished from an EGFR inhibitor effect.

3.9.3 Sensitivity of xenografted LIM1215 cells after treatment with 5-Azacytidine

In LIM1215 cells, sensitivity towards Erlotinib and Gefitinib was increased by Valproat. Valproat also led to an increase of *AREG* or *EREG* expression (see figure 6, figure 7 and figure 10). As a next step, it was tested if the sensitivity increase could also be observed *in vivo*. In cooperation with Maria Rivera (Experimental pharmacology, Max-Delbrück Center Berlin-Buch), LIM1215 cells were processed as mouse xenografts to evaluate the influence of Valproat on tumor growth with and without Erlotinib treatment (for experimental details please refer to section 2.12.2). In addition, 5-Azacytidine, a DNA methyltransferase whose effects in mice is well characterized and which has similar but not identical functions as DAC, was tested for its influence on the tumor growth in presence or absence of Erlotinib. The results of three independent experiments are shown in figure 54.

Figure 54:

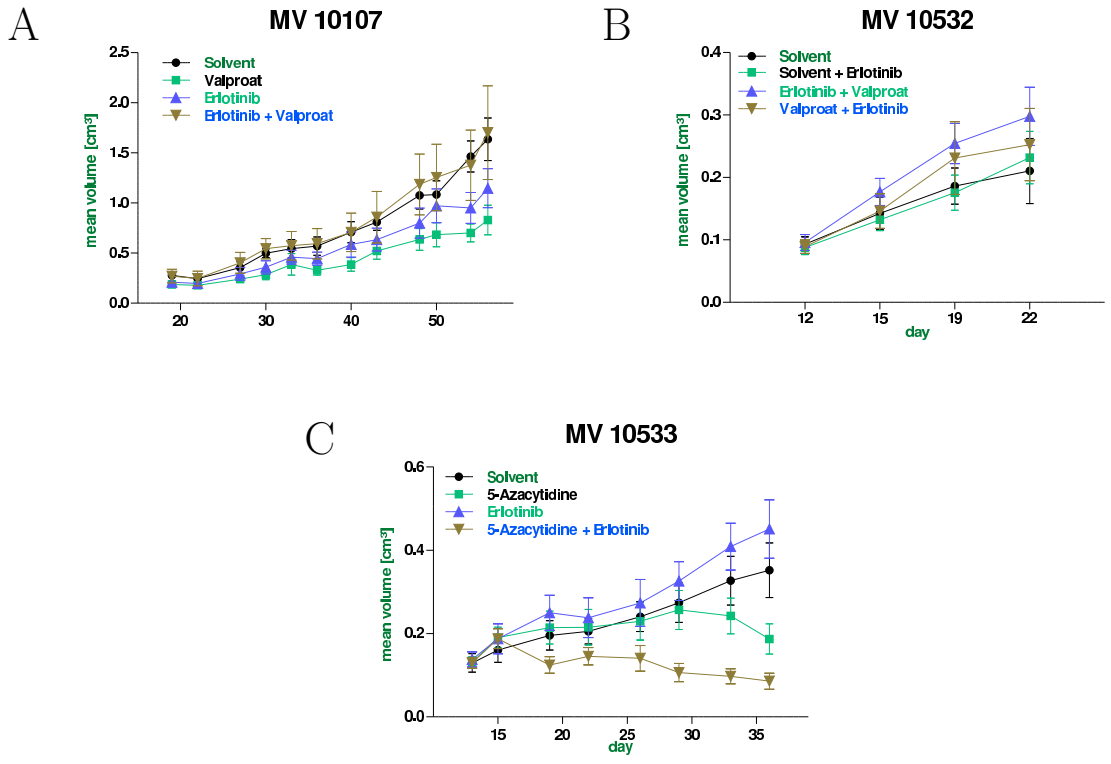


Figure 54: Experiments with xenografted LIM1215 cells: LIM1215 cells were processed as xenografts and mice were subsequently treated as described in section 2.12.2. Three independent experiments are shown. Each data curve represents the mean volume of 5 mice per treatment with the error bars representing the standard deviations. The tumor volumes were determined two times per week as described in section 2.12.2.

In the mouse experiment MV10107 it was examined, if Valproat could lead to a higher sensitivity towards Erlotinib treatment when applied in combination with Erlotinib (see figure 54 A). Valproat as well as Erlotinib caused smaller tumor volumes

compared to the xenografts in solvent treated mice. However, the volumes were not significantly smaller. But, when treating Valproat in combination with Erlotinib the xenografts grew at similar rates as the xenografts of solvent treated mice. Because the differences were not significant the experiment was repeated. In experiment MV10532, treatment was also adapted to the treatment modes used in cell culture, which means Valproat treatment was done prior to Erlotinib treatment. Additionally, in control mice, Valproat treatment was done after Erlotinib treatment to investigate if the treatment mode influences the outcome of the experiment. As seen in figure 54 B, the xenografts of the solvent treated mice grew similar like the xenografts of the Erlotinib treated mice. Interestingly, both combination approaches, Valproat prior or after Erlotinib treatment, caused increased tumor volumes compared to solvent treatment. The reason remains unknown. Summing up, Valproat did not increase sensitivity towards Erlotinib *in vivo*. In the experiment MV10533, it was tested, if the DNA methyltransferase inhibitor 5-Azacytidine influences sensitivity towards Erlotinib *in vivo*. As described in figure 54 C, the tumor volumes of Erlotinib treated mice did not change significantly compared to solvent treated mice. Also 5-Azacytidine did not affect tumor volumes until day 29. However, a strong reduction of the tumor volumes was observed in the xenografts of the mice treated with a combination of 5-Azacytidine and Erlotinib. Therefore, 5-Azacytidine increased sensitivity towards Erlotinib *in vivo*.

After the experiment, *AREG* and *EREG* mRNA expression were tested. As seen in figure 55, treatment with Valproat or with 5-Azacytidine did not lead to an increased *AREG* or *EREG* mRNA expression in the xenografts compared to the xenografts of the solvent treated mice. The ΔCt-values in the *AREG* gene expression analysis were similar in all experiments for all samples between -2 and -4, and the ΔCT-values in the *EREG* gene expression analysis ranged in all samples in all experiments between -6 and -8. The values were equal to the ΔCt-values obtained for untreated LIM1215 cells in cell culture experiments. In addition, mouse *AREG* and *EREG* mRNA expression were tested. But, neither mouse *AREG* nor mouse *EREG* mRNA expression were increased after Valproat or 5-Azacytidine treatment. After all, no xenograft sample showed any expression of mouse *AREG* or *EREG* mRNA (data not shown).

Figure 55:

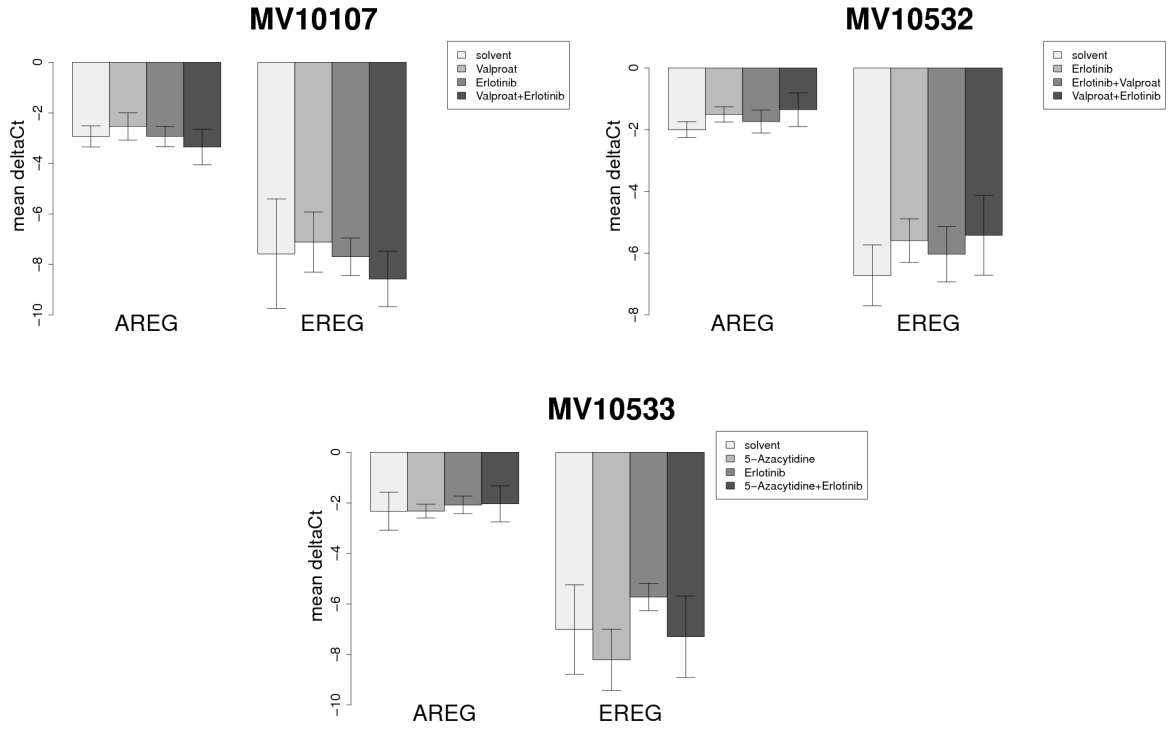


Figure 55: *AREG* and *EREG* mRNA expression in xenografts: human AREG and *EREG* mRNA expression of the xenografts obtained in the mouse experiments were measured by real-time PCR as described in section 2.6. The ΔCt-values were calculated by subtraction of the AREG or EREG Ct-value from the control Ct-value For each experiment the mean ΔCt values of five samples per treatment are shown together with the standard deviations.

Additionally, the tumors were fixed with formalin and embedded in paraffin by a standard protocol performed in the in-house facility of the Institute for Pathology. The resulting paraffin-blocks were cut in 3 µm slices and transferred to glass slides. Immunohistochemistry experiments (IHC) were performed to detect the AREG protein (see section 2.11). The experiment was analyzed together with Dr. Florian Rossner (Institute for Pathology, Charité, Berlin) The results are shown in table 21

Table 21: IHC evaluation of AREG protein expression in the mouse xenograft experiments: fail: sample could not be analyzed, -: no AREG protein, +/-: some cells show positive, some show negative AREG protein expression, +: cells show positive AREG protein expression, ++: cells show strong AREG protein expression.

experiment	treatment	fail	-	-/+	+	++
MV10107	Solvent		2	2	1	
	Valproat		1	2	1	1
	Erlotinib		2	1	1	1
	Erlotinib + Valproat		2	2		1
MV10532	Solvent	2	2			1
	Solvent + Erlotinib	4	1			
	Erlotinib + Valproat	1	1	1	1	1
	Valproat + Erlotinib	2		1	2	
MV10533	Solvent				4	1
	5-Azacytidine	2			2	1
	Erlotinib	1		1	1	2
	5-Azacytidine + Erlotinib	5				

In all three experiments the expression of AREG protein varied strongly among the different treatment modes. No correlation is visible between the AREG protein expression and the treatment. High failure rates of *AREG* IHC were also observed in some examples (see table 21). In most of the cases this was due to too small sample sizes to perform IHC. The sample size depended on the tumor volume after treatment. Nevertheless, an example, on how the samples looked like is shown in figure 56. Here microscopy pictures are shown representing the results of MV10107. For each treatment a sample is given showing no AREG protein, a sample showing intermediate AREG protein and a sample showing high AREG protein expression. To summarize, the expression of *AREG* was not consistently changed by different treatment options *in vivo*.

Figure 56:

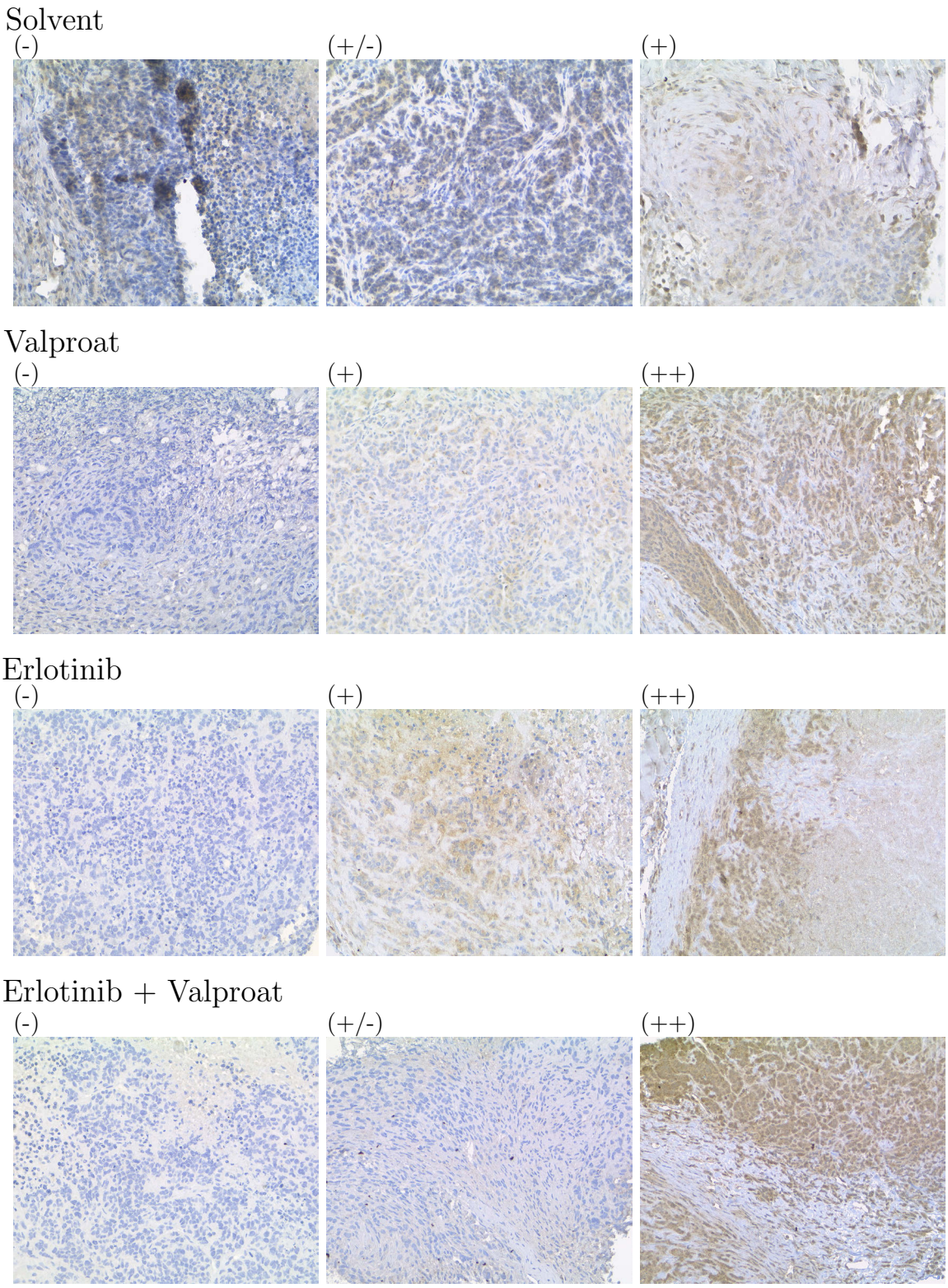


Figure 56: AREG protein expression in xenografted LIM1215 cells of experiment MV10107: The AREG protein expression was evaluated by IHC. For each treatment option three figures are shown representing three samples with different AREG protein expression. Magnification factor: 200x, brown: AREG protein staining, blue: nuclei staining. (-): no AREG protein, (+/-): some cells show positive, some show negative AREG protein expression, (+): cells show positive AREG protein expression, (++): cells show strong AREG protein expression.

Figure 57:

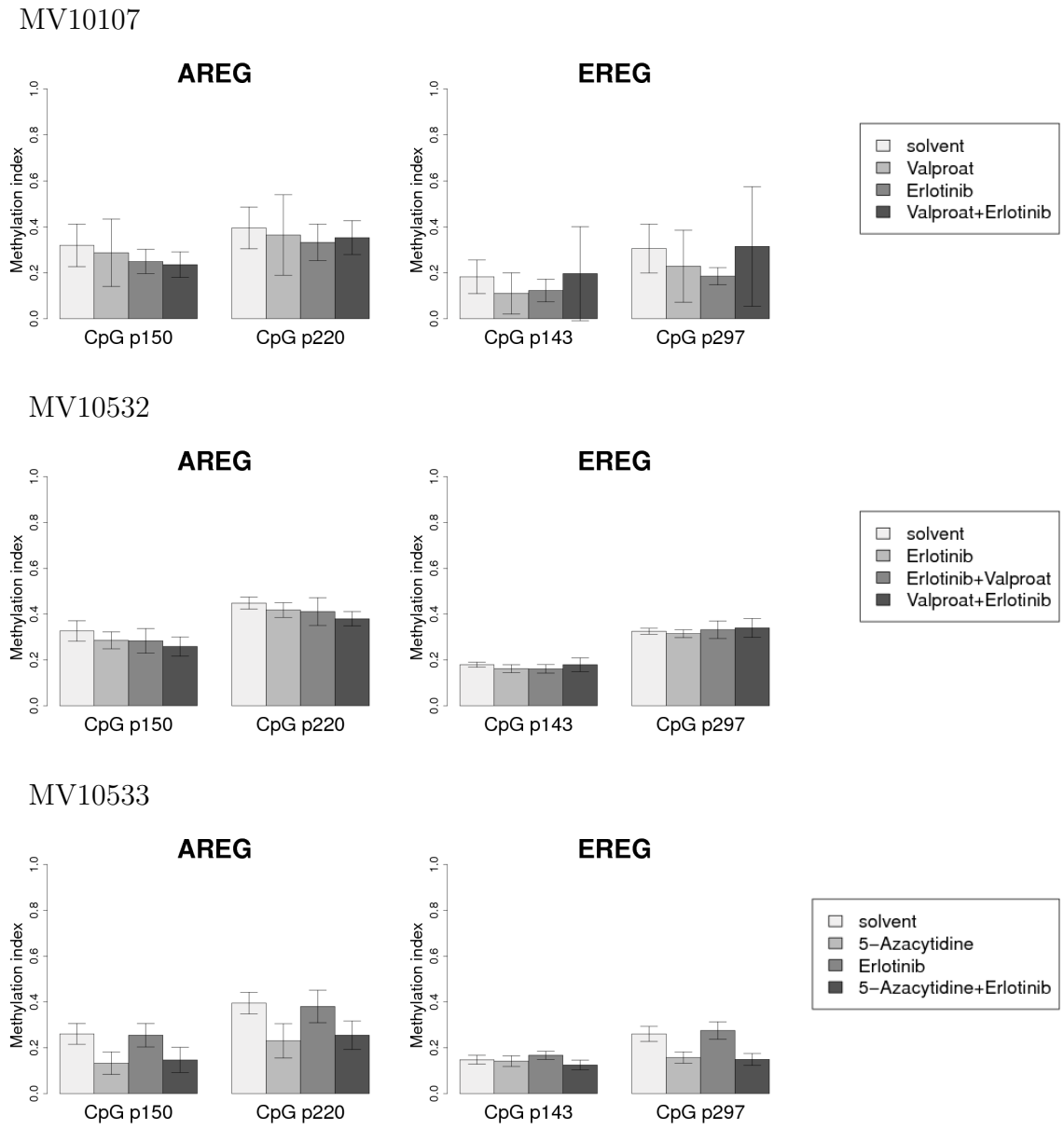


Figure 57: Methylation of CpGs within the *AREG* or *EREG* genes in xenografts: The methylation of *AREG* CpG p150 and p220 as well as *EREG* CpG p143 and p297 was measured in all xenografts obtained in the mouse experiments as described in section 2.12.1. The methylation index shows the content of methylation indicating products within the primer extension products in msSNuPE-experiments and is an indicator for DNA methylation at the given CpG. A methylation index of 1 means fully methylated whereas a methylation index of 0 means fully unmethylated.

Methylation of the same *AREG* and *EREG* intragenic CpGs, which were tested in cell culture, were also tested in each xenograft (see figure 57). In comparison to the LIM1215 cell line in cell culture, the xenografted LIM1215 cells showed a treatment independent reduction of methylation of *AREG* CpG p150 and CpG p220. In cell culture the CpG methylation of the *AREG* intragenic CpGs ranged from 0.4 to 0.6, whereas in the xenografts the methylation index ranged from 0.2 to 0.4 (compare figures 14 and 57).

In comparison the methylation of *EREG* CpGs p142 and p297 did not differ strongly between cell culture and xenografted LIM1215 cells. In both cases, the methylation index of the tested *EREG* intragenic CpGs ranged from 0.2 to 0.4. Similar to the cell culture results, Valproat treatment did not influence methlylation of the tested CpGs (see figures 18 and 19). Methylation of *AREG* CpG p150, *AREG* CpG p220 and *EREG* CpG p297 was affected by 5-Azacytidine when the mice were treated with 5-Azacytidine alone or in combination with Erlotinib. That means similar to the cell culture results, a methylatransferase inhibitor led to a demethylation of the tested GpGs *in vivo* (see figures 15 and 16).

3.10 Over-expression of *AREG* in LIM1215 has no significant effect on sensitivity towards EGFR inhibitors

Figure 58:

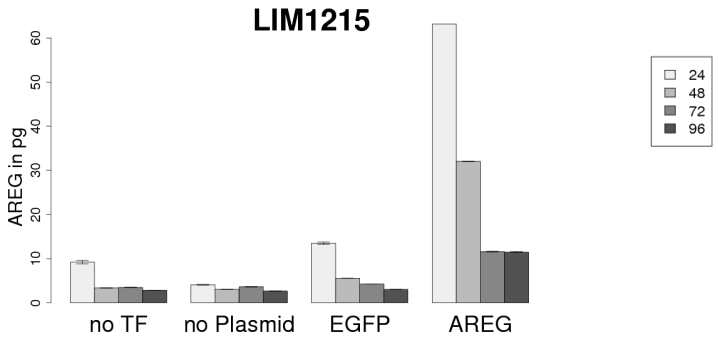


Figure 58: AREG protein expression after transient transfection: LIM1215 cells were transfected transiently with the Plasmid EX-A0114-M02 containing the *AREG* sequence (GeneCopoeia) and a control plasmid containing the *EGFP* sequence (Invitrogen). AREG protein expression was determined by ELISA and was normalized per 1 µg total protein.

As described in section 3.2.2 the mRNA expression and protein levels of AREG increased by up to 4-fold upon treatment of LIM1215 cells with Valproat. At the same time the sensitivity of these cells towards the EGFR inhibitors Erlotinib and Gefitinib was increased following Valproat treatment (see section 3.9.2 and figure 47). These observations raised the question to what extent AREG is functionally involved in the increased response at LIM1215 cells towards EGFR inhibitors.

LIM1215 cells were transfected transiently with an *AREG* expression plasmid (EX-A0114-M02) and AREG protein expression was monitored by ELISA 24 h to 96 h post transfection. 24 h after transfection, AREG protein levels were up to 6-fold higher compared to the untransfected LIM1215 cells (see figure 58). However, protein levels decreased down to approximately 3-fold 96 h after transfection. AREG levels also increased after 24 h in LIM1215 cells transfected with an *EGFP* control plasmid, but only up to 1.5-fold.

To test whether AREG levels correlate with the sensitivity towards EGFR inhibitors, LIM1215 cells were transfected either with the *AREG* expression plasmid EX-A0114-M02 or the *EGFP* control plasmid. Cells were seeded into 96-well plates and XTT proliferation experiments were performed, evaluating cell growth after Erlotinib and Gefitinib treatment.

Figure 59:

LIM1215:

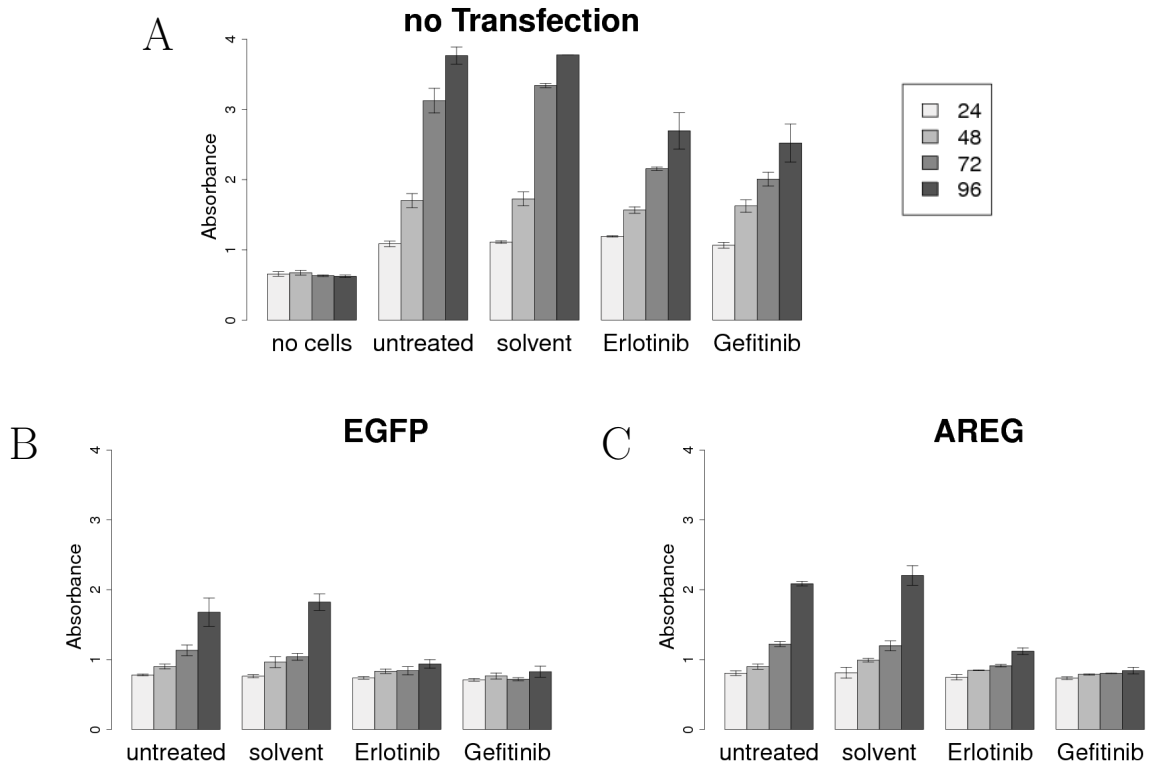


Figure 59: XTT proliferation after transient Transfection: LIM1215 cells were transfected transiently with the *AREG* sequence containing Plasmid EX-A0114-M02 or the *EGFP* control plasmid. 24 h after transfection cells were plated into 96-well plates and XTT experiments were performed as described earlier to evaluate the sensitivity towards Erlotinib and Gefitinib.

As shown in figure 59 non-transfected LIM1215 cells behaved similar as determined before. After transfection, both *EGFP* and *AREG* transfected cells grew much slower. In contrast to the non-transfected cells, transfected LIM1215 cells even stopped growing after treatment with Erlotinib and Gefitinib (compare figure 59 A with B and C). The XTT absorbance of the EGFR inhibitor treated non-transfected cells was approximately 25 % lower at 72 h and 96 h than the XTT absorbance of the untreated and solvent treated control cells (see figure 59 A). Interestingly, the XTT absorbance of the EGFR inhibitor treated transfected cells (transfection plasmid independent) is at 96 h approximately 50 % lower than the untreated and solvent treated control cells (see figure 59 B and C). However, the transfection method itself influences growth rates negatively, too. As mentioned before, it was observed that transfecting the EGFP control plasmid

causes also an increase of AREG protein levels (see figure 58). To verify this result, *AREG* mRNA expression was measured after the XTT proliferation experiment. In figure 60 is shown that *AREG* mRNA was also increased after transfecting the EGFP plasmid. That means, AREG might be responsible for the increase of sensitivity towards EGFR inhibitors, but since transient transfected cells grew much slower than non-transfected cells a second approach was necessary to clarify the influence of AREG onto EGFR inhibitor sensitivity.

Figure 60:

LIM1215:

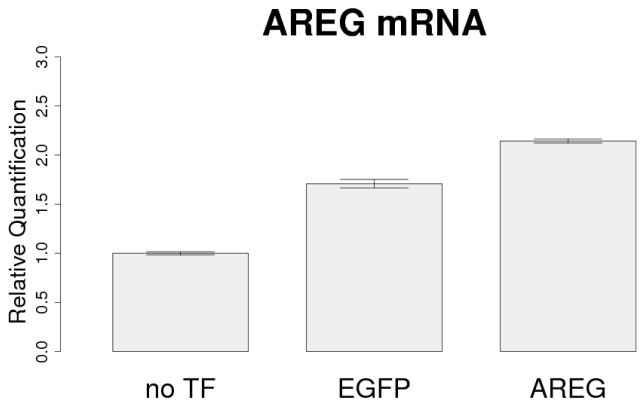


Figure 60: *AREG* mRNA expression after the XTT proliferation experiment: LIM1215 cells were transfected transiently with the plasmid EX-A0114-M02 containing the *AREG* sequence and the *EGFP* control plasmid. 24 h after transfection cells were plated into a 10 cm petri dish and incubated until the XTT experiment was finished (see figure 59). After this, RNA was isolated and *AREG* mRNA expression was detected by real-time PCR.

Lentiviruses were produced as described in section 2.9.2. LIM1215 cells were infected with lenitviruses containing either a plasmid with the complete *AREG* coding sequence (AREG), an *AREG* sequence lacking all C-terminal aminoacids encoded by *AREG* exon 5 (AREG-short or A-short) or an empty control vector (empty). Infected cells were tested for *AREG* mRNA expression and protein levels (figure 61). Real-time PCR analysis (TaqMan®) showed that the *AREG* mRNA expression was increased up to 3-fold in the "AREG" vector infected cells. No increase was observed for the "A-short" vector infected and the "empty" vector infected cells. The *AREG*-Taqman assay used detects the *AREG* cDNA sequence with primers spanning exon 4 and 5. The sequence of "A-short" ends at exon 4. Therefore, "A-short" could not be detected by this assay. To circumvent the problem, mRNA expression was determined by SybrGreen® using primers binding in exon 1 and 2. By this approach, it could be demonstrated that "A-short" as well as "AREG" infected cells had an increased *AREG* mRNA expression. SybrGreen® was performed twice. *AREG* mRNA expression was increased by 3 to 4-fold (AREG) or 5 to 7 -fold (A-short). Also an increased AREG protein level in the supernatant could be demonstrated for the "AREG" and "A-short" infected cells (figure

61, right). In contrast to the *AREG* mRNA expression in the transfected cells, protein levels increased only 3 to 3.5-fold. However, since Valproat treatment also led to an increase of AREG protein by 2 to 4 fold (see figure 7) and *AREG* mRNA expression by 2 to 4 fold (see figure 6), LIM1215 cell growth and sensitivity towards Erlotinib and Gefitinib was tested with the infected cells as described before.

Figure 61:

LIM1215:

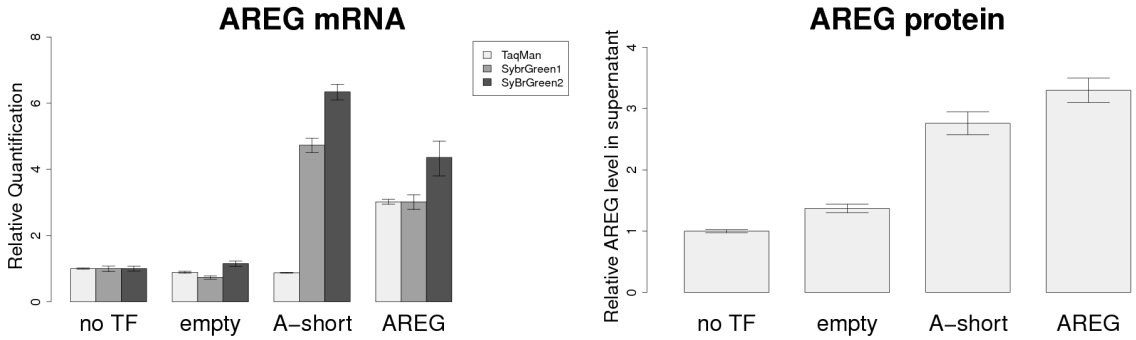


Figure 61: *AREG* expression after lentiviral transfection: LIM1215 cells were infected by lentiviruses containing *AREG* gene expression plasmids (AREG), *AREG* gene w/o C-terminus expression plasmids (A-short) or the empty plasmids. Left: *AREG* mRNA expression was detected by TaqMan® or SybrGreen® real-time PCR. Relative Quantification was calculated according to the non-transfected control. Right: The AREG protein amount in the supernatant was detected by ELISA and normalized to the corresponding RNA concentration. Here, the relative AREG level compared to the untransfected control is shown.

XTT absorbances dropped in non-transfected LIM1215 cells by approximately 20 to 25 % after 48, 72 and 96 h of Erlotinib and Gefitinib treatment compared to the untreated or solvent treated cells (see figure 62 A). Therefore, the sensitivity towards Erlotinib and Gefitinib was similar as seen before (see 42 A). No change in sensitivity was observed, when infecting LIM1215 cells with the plasmids "AREG" or "AREG-short" (see figure 62 C-D). To conclude, AREG over-expression, which was achieved by lentiviral infection, did not lead to an increase of sensitivity towards the EGFR inhibitors.

Figure 62:

LIM1215:

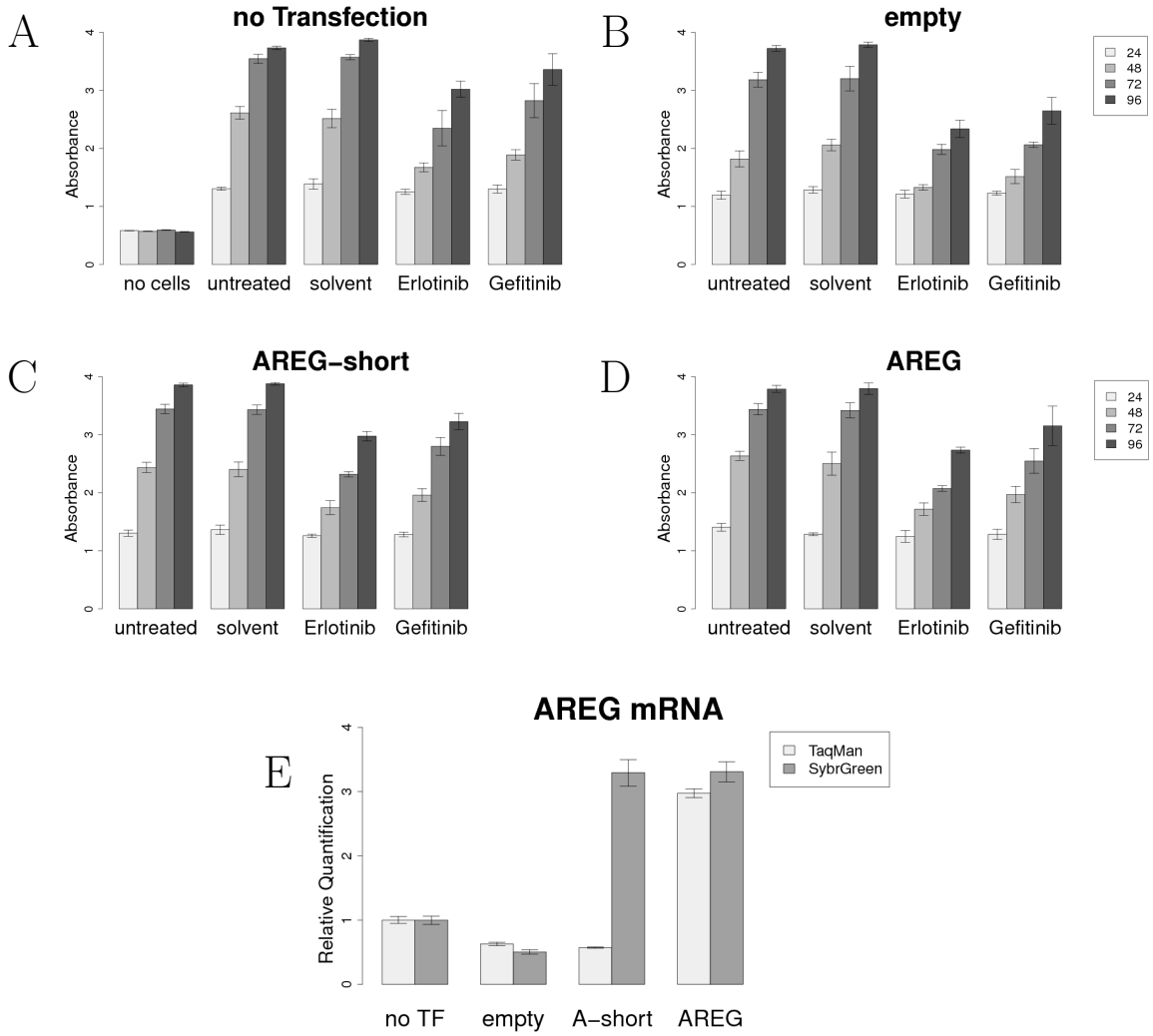


Figure 62: XTT experiment after lentiviral infection: LIM1215 cells were either used directly (A) or they were infected by lentiviruses to introduce an empty plasmid (B), an *AREG* plasmid w/o C-terminus (AREG-short, C) or an an *AREG*-expressing plasmid (D). After determination of the *AREG* expression (see figure 61), cells were seeded into 96-well plates. Sensitivity towards EGFR inhibitors was determined by XTT experiments as described earlier. XTT-absorbance was measured 24 h after XTT solution dispense. Simultaneously, the cells were seeded into 10 cm dishes for RNA isolation after the experiment. *AREG* mRNA expression was tested again by real-time PCR (E).

4 Discussion

4.1 The EGFR-ligands are regulated by versatile mechanisms

AREG and other EGFR ligand genes such as *TGF-α*, *HB-EGF*, *BTC* and *EREG* can be activated by different mechanisms. One example is the activation via EGFR-signaling by an autocrine loop. In a study using prostate stromal cells, Soerensen *et al.* found out that *AREG*, but also *HB-EGF* expression increased after stimulation of EGFR with EGF or HB-EGF,[247] indicating that HB-EGF can activate its own gene. Examinations in zebrafish verified the results for HB-EGF, but in contrast, *BTC* was observed to be down-regulated upon BTC-treatment.[248] The underlying mechanism of the autocrine regulation includes downstream signaling pathways like the MAPK- and the PI3K-pathway. A connection between *AREG* gene expression and MAPK-pathway is strengthened by the fact that during puberty and pregnancy *AREG* gene expression is stimulated by estrogen and progesterone.[31,32] It was shown that both receptors can activate MAPK-signaling.[249] In addition, Sizemore and Cox showed that farnesyltransferase inhibitors caused a decreased AREG and TGF-α expression, hinting for RAS as a protein involved in AREG regulation.[250] In cancer cells, the MAPK and PI3K-pathways might not only be activated via receptors, but also via activating mutations. As described in section 1.2.3 *KRAS*, *BRAF* and *PI3K* belong to the most frequently mutated genes within these pathways in colorectal cancer. Therefore, it was analyzed in this work if mutations in these genes correlate with *AREG* or *EREG* expression. Although *AREG* and *EREG* were differentially expressed in the cell lines tested (see figure 3), the expression did not correlate with mutations within the MAPK- or PI3K-pathway. SW480 as well as HCT116 cells harbor *KRAS* mutations (see table 12). However, SW480 was a low *AREG* and *EREG* expressing cell line whereas HCT116 was a high *AREG* and *EREG* expressing cell line (see figure 3). Also *BRAF* mutated cells were found among the low *AREG* and *EREG* expressing cells (RKO) and among the high *AREG* and *EREG* expressing cells (HT29), too. RKO and HT29 cells are also examples for *PIK3CA* mutated cells harboring a deregulated PI3K-pathway. If the mutation states of MAPK- or PI3K-signaling components influence the *AREG* and *EREG* expression, as expected from literature, they may not be responsible alone for the *AREG* and *EREG* regulation within the cell lines used in this work.

Another possible mechanism for *AREG* regulation is the regulation via activated wnt-signaling. TCF/LEF binding sites were identified within the *AREG* promoter.[251] Therefore, the cells were analysed for *APC* mutations, which deregulate the wnt-signalling. Similar to *KRAS* and *BRAF*, mutations within the *APC* gene were found in high *AREG* expressing cell lines (CaCO2,[252] HT29 (COSMIC Cell Lines Project)) as well as in low *AREG* expressing cell lines (SW480,[252] Colo678 (COSMIC Cell Lines Project)) Also, cells with wildtype-*APC* are within the low *AREG* expressing cells (RKO (COSMIC Cell Lines Project)) and within the high *AREG* expressing cells (HCT116[252]).

Nevertheless, the impact of wnt-signaling onto *AREG* expression cannot be deduced from the analysis of *APC* mutations within the cell lines alone because publications also suggest that wnt-signaling might also be activated by β-catenin mutations.[253]

A further regulatory mechanism of different EGFR-ligand genes is cAMP-mediated regulation. In a study using murine granulosa cells *AREG*, *EREG* and *BTC*-mRNA expression was increased in the presence of cAMP.[254] Also a CRE-element was identified within the *AREG* promoter[255,256] supporting theses findings. The CRE-element serves as a binding site for the CRE-binding protein (CREB), which recruits the CREB-binding protein (CBP)/p300 coactivator complex. This complex is then able to interact with transcription factors TFIIB and TFIID as well as with the RNA polymerase II to initiate transcription.[257,258] CBP and p300 are histone acetyltransferases (HATs).[259] Thus, if cAMP-mediated regulation takes place, CBP/p300 could contribute to the *AREG* expression via an open chromatin structure at the *AREG* promoter. Because HDACis also shift the chromatin structure to open chromatin, this mechanism could serve as a hint, why *AREG* and *EREG* expression increases after HDACi treatment and suggest that the *AREG* promoter is involved in *AREG* regulation.

In addition to a HAT-mediated open chromatin structure at the *AREG* promoter, DNA methylation is another possible epigenetic regulatory mechanism influencing *AREG*. As already described in section 1.4.2, chromatin structure and DNA methylation are tightly linked to each other. For *EREG*, a methylation-dependent regulation was already suggested, since *EREG* promoter methylation was observed in gastric cancer cells.[260] Also, experiments comparing the global mRNA expression between HCT116 and the *DNMT1/3b* lacking cell line HCT116^{DKO} hint for a DNA methylation-dependent regulation of *AREG*, *EREG* and *BTC*.[261,240] (Sers *et al.* 2009, Gius *et al.* 2004 and data not shown) Following these data, a deeper insight into the epigenetic regulation of *AREG* and to a lower extent *EREG* was made in this work. The results are discussed in section 4.2.

In addition to the multiple mechanisms acting at the genes' levels, AREG and EREG proteins can also be regulated by indirect mechanisms. The ADAM-17 protein for instance is responsible for the segregation of several EGFR-ligands into the extracellular space. A decreased expression of TIMP3, an ADAM-17 inhibitor, was shown e.g. to correlate with an increased release of AREG, HB-EGF and TGF-α.[262]

To summarize, no clear connection could be made between distinct signaling pathways and *AREG* and *EREG* expression in colon cancer cell lines. The mutations of genes within these pathways described in literature to influence *AREG* and *EREG* regulation did not correlate with the *AREG* and *EREG* expression in this study.

4.2 *AREG* and *EREG* are regulated via epigenetic mechanisms

Due to the availability of new genome-wide detection methods as for example Second Generation Bisulfite Sequencing, it was found that a major part of DNA methylation is not located within CpG-islands of promoter regions. They were rather detected in regions with low CpG-density.[80] These CpGs might influence gene expression. It was e.g. shown in a genome wide study that exon methylation is linked to transcriptional silencing.[79] Also in a different study using the human B-cell methylome, it was shown that although only a subset of all CpG-islands are methylated (about 26 %), the ratio of CpG-islands being methylated within genes is higher (approximately 36 %).[263] By comparing their data with gene expression data, the authors also showed that increased intragenic CpG-island methylation is correlated with reduced expression of the corresponding genes. These studies emphasize that intragenic methylation patterns might become important issues in future research, since they might be responsible for regulatory mechanisms, which are not yet completely understood. Nevertheless, the data presented in this work might contribute strongly to an uncovering of these regulatory mechanisms.

As described in section 3.1, *AREG* was differentially expressed within the cell lines tested. However, the expression did not correlate with promoter methylation but it correlated with the methylation of CpGs within the gene (CpG p150, CpG p220, see section 3.5). The results were supported by experiments with the DNMT- inhibitor DAC, because demethylation of the intragenic CpGs was observed in combination with an increase in *AREG* mRNA and protein expression (see figures 4, 5 and 15). Furthermore, the increase in *AREG* expression was higher in cells with lower initial *AREG* expression and higher *AREG* intragene methylation. It was lower in cells with higher initial *AREG* expression and lower *AREG* intragene methylation. These facts suggested a direct methylation-dependent mechanism of *AREG* gene expression, which is different to promoter methylation-dependent gene regulation mechanisms.

With the exception of RKO cells, *EREG* promoter methylation did not correlate with gene expression (see figure 12). RKO cells showed high methylation within the *EREG* promoter sequence and they also showed very low *EREG* gene expression (see figure 3). Therefore, in RKO cells a promoter methylation-dependent regulatory mechanism of the *EREG* gene might be active, like it was already described in gastric cancer cells.[260] *EREG* promoter methylation might also be due to the CpG-island-methylator phenotype. RKO is the only cell line within this work, which was described to be CIMP-positive.[264] However, the fact that SW480 cells showed no *EREG* promoter methylation but a similar low *EREG* gene expression suggests that there are also other regulatory mechanisms responsible for *EREG* expression. One possible mechanism might be an intra-gene methylation-dependent regulatory mechanism, as it was seen for *AREG*. But in contrast to *AREG*, the methylation of the CpGs investigated within the *EREG* gene

did not correlate with the observed *EREG* gene expression (see figures 3 and 14). Nevertheless, similar to *AREG*, DAC treatment led to a high *EREG* increase in low *EREG* expressing cells and to a low *EREG* increase in high *EREG* expressing cells (see figure 4). But changes in CpG-methylation did not correlate with changes in expression. First, no demethylation of the intragenic CpGs was observed in the cell lines SW480 and LIM1215, although *EREG* mRNA expression increased upon treatment; and second, although the highest demethylation of the *EREG* intragenic CpGs was observed in the CaCO2 cell line, DAC did not result in an increase of *EREG* expression (compare figure 16 and figure 4). The DAC-treated SW480 and LIM1215 cells showed a demethylation of *AREG* CpG p150 and CpG p220 (see figure 16), indicating that the missing effect on *EREG* CpG p143 and CpG p297, was not caused by a failure in treatment. Probably the regulation of *EREG* is similar to *AREG*, but the tested CpGs within the *EREG* gene are not responsible for this regulation. Even though the results lead to a methylation-dependent regulation, it should be noted that DAC, as well as all other DNMT inhibitors, have a global effect on DNA methylation. Therefore, it cannot be excluded that changes in *EREG* expression by DAC might also be mediated by demethylation-dependent expression changes in other genes. *AREG* might be a candidate, since autocrine activation of the MAPK- and PI3K-pathways was shown to influence also *EREG* expression.[254] In the same way, *AREG* might also influence its own expression.

Histone deacetylase inhibitor treatment (HDACis) also led to an up-regulation of *AREG* and *EREG* gene expression. While Valproat mainly had effects on LIM1215, SW480 and RKO cells (see figures 6 and 7), TSA affected all cell lines (see figures 8 and 9). During his master thesis, Stephan Bartels (Charité Berlin, Institute for Pathology) aimed at defining the mechanism how HDACi treatment affects the MAPK- and PI3K-pathways. He observed in three cell lines that Valproat treatment led to an increase in phosphorylation of MEK, ERK and AKT, indicating an activation of these pathways.[241] Since the direct mechanistic processes were not addressed, yet, it remains unclear, to what extent MAPK and PI3K-pathway signals and epigenetic regulatory mechanisms influence *AREG* and *EREG* expression upon HDACi treatment. According to current knowledge, three regulatory mechanisms might be possible:
1. HDACi treatment might influence *AREG* and/or *EREG* expression by shifting the chromatin structure at the genes' promoters to open chromatin. AREG and EREG then activate the MAPK- and PI3K-pathways via an autocrine loop. 2. HDACi treatment might influence the MAPK- and PI3K-pathway independent of *AREG* or *EREG*. *AREG* and *EREG* are then activated as target genes of the pathways. 3. *AREG* and *EREG* expression are influenced by both, pathway-signal dependent mechanisms and epigenetic mechanisms.

Besides the arguments showing a possible MAPK- and PI3K- pathway dependent regulatory mechanism (see section 4.1), the following argument strengthens an epigenetic mechanism influencing *AREG* or *EREG* expression upon HDACi treatment: As

already described in section 1.4.2, DNA methylation is tightly linked to chromatin structure. Recruitment of HDACs to methylated DNA via MBD-containing proteins[101] leads to transcriptional silencing which can be relieved via HDACis like TSA.[102,101,103] By this mechanism, *AREG* gene activation by HDACi-induced *AREG* promoter acetylation might be possible, which might be similar to the effect after HAT-recruitment in the cAMP-mediated regulatory mechanism mentioned above.

In addition to the possible effects on signaling mechanisms or epigenetic mechanism other mechanisms, not directly impinging on the *AREG* promoter, might also be responsible for the increased *AREG* and *EREG* expression upon HDACi treatment. 1. Global deacetylation influences the intra-cellular pH by changing the amount of free acetate anions which non-specifically influence the effect of the HDACis on gene expression.[265] 2. Side effects of Valproat itself might influence expression. It was e.g. described that Valproat has effects on processes like inhibition of GABA-transaminase[266] or reduction of Inositol-3-phosphate levels.[267] Inositol-3-phosphate is a second messenger, which e.g. changes ion concentrations within the cell. This might also affect other genes.

To summarize, there is great evidence that *AREG* and *EREG* expression are regulated by epigenetic mechanisms. Most likely, this is not the only mechanism regulating *AREG* and *EREG* gene expression. Probably, there is a complex interplay between signaling-mediated and epigenetically determined processes controlling the expression of both genes.

4.3 *AREG* expression is influenced by zinc finger associated transcription factors

As described in section 3.7, *CTCF* was highly expressed in all cell lines, whereas *ZBTB33* and *ZBTB4* were differentially expressed. The expression of the latter transcription factor genes was inversely correlated with the *AREG* gene expression (see figure 32). An argument showing that *AREG* and *ZBTB33* expression are also correlated functionally was given by the PhD-student Natalia Kuhn (Charité Berlin, Institute for Pathology), who used siRNAs targeting *ZBTB33*. In these experiments it was shown that an siRNA-mediated decreased expression of *ZBTB33* was accompanied with an increased *AREG* expression (see figure 33). Furthermore, bioinformatic approaches searching for potential binding sites showed that ZBTB33 but also CTCF might be able to bind within the *AREG* gene. CTCF-binding sites were identified by Sascha Tierling (University of Saarbrücken) using the webtool "CTCF-binding sites" (http://bsproteomics.essex.ac.uk:8080). ZBTB33-binding sites were identified by Juliane Perner (Max-Planck-Institute for Molecular Genetics, Berlin) based on a recent publication of Blattler *et al.*[109] (see figure 63). Interestingly, one CTCF- and one ZBTB33-binding site are located within *AREG* exon 2. The ZBTB33-binding site is located only two bases downstream the *AREG* CpG p150 in reverse complement

orientation and it is the only site, which is located directly on the *AREG* gene (see figures 63 B and C). The next ZBTB33-binding sites are either located approximately 9.5 kb upstream the *AREG* promoter or approximately 15 kb downstream the last *AREG* exon. *EREG* contained several ZBTB33-binding sites, too, which might hint for a similar regulatory mechanism (see figure 63 A). In the publication of Blattler *et al.*, ENCODE data were used to identify interaction partners of ZBTB33.[109,110] They showed that ZBTB33 is associated with RNA-polymerase II and other transcription factors. Therefore, the ZBTB33-binding site, identified within *AREG* exon 2, might hint for *AREG* exon 2 to have promoter function.

To prove that, several *Firefly-luciferase*-gene containig reporter plasmids were generated and promoter function analysis experiments were performed (see section 3.6). The *AREG* CpG p150 and CpG p220 containing exon 2 was used for analysis alone or coupled to the *AREG* promoter sequence or the Tata-binding motif, mimicking a generic promoter. The complete *AREG* exon 2 sequence, as well as the exon 2 parts containing only CpG p150 or CpG p220, did not show promoter function in any of the cell lines tested when analyzed alone (see figures 25 A and 26 A). However, in HCT116 cells, the reverse-oriented exon 2 sequence showed promoter function (see figures 30 A). Also, when coupling the *AREG* exon 2 sequence in reverse orientation to a Tata-binding motif, promoter function of the Tata motif increased highly significant in HCT116 cells (up to 10-fold), but also in the cell lines LIM1215 and CaCO2 (see figures 25 D, 26 D, 28 C-D and 29 C). In HCT116 cells promoter function increased even more, when the reporter plasmid was methylated *in vitro* (see figure 28 C). These data suggest a) a promoter function of the *AREG* exon 2 sequence which is b) methylation- and orientation-dependent. A second interesting result was observed, when coupling the *AREG* exon 2 sequence in both orientations to the *AREG* promoter. The promoter function compared to the *AREG* promoter alone was significantly increased in HCT116 cells (see figures 26 B and 28 A and B). It also increased slightly in LIM1215 cells (see figure 25 B). However, a reduction of promoter function was observed in CaCO2 cells (see figure 29 B), which was significant for the reverse-oriented exon 2 sequence. Therefore, a cell type specific regulatory mechanism might occur, which is sequence orientation-independent. However, it is unclear why promoter function dropped in CaCO2 cells to a lower value than the promoter function of the *AREG* promoter alone.

As already described, ZBTB33 might be responsible for the cell type specific promoter function of *AREG* exon 2, because it is higher expressed in LIM1215 and HCT116 cells than in CaCO2 cells (see figure 32). Nevertheless, ZBTB33 might not be alone responsible, because an orientation-dependent promoter function of *AREG* exon 2 in combination with the Tata-binding motif was seen in CaCO2 cells, too. To clarify the influence of ZBTB33 on the promoter function of *AREG* exon 2, experiments are currently in progress to mutate the identified ZBTB33-binding site within the *AREG* exon 2 sequence and afterwards test the constructs in promoter analysis experiments.

Experiments addressing the influence of CTCF, the second candidate which might be involved in *AREG* expression, were already performed in HCT116 cells. Interestingly, mutation of the CTCF-binding site in the *AREG* exon 2 coupled in forward orientation to the *AREG* promoter led to a reduction of promoter function compared to the *AREG* promoter alone (see figure 30 B). Remarkably, a complete loss of promoter function was observed, when the mutated *AREG* exon 2 was coupled in reverse-orientation to the *AREG* promoter. These data suggest that CTCF influences the promoter function of the *AREG* exon 2 when coupled to the *AREG* promoter. But, in contrast to *ZBTB33*, *CTCF* mRNA expression was comparably high in all cells tested (see figure 32). Data obtained by Defossez *et al.* support a hypothesis including both CTCF and ZBTB33 in *AREG* regulation, because they showed that CTCF can be an interaction partner of ZBTB33, when there is a ZBTB33-binding site in close proximity to a CTCF-binding site.[268] The authors also showed that ZBTB33 inhibits the insulator function of CTCF by using an insulation assay *in vitro*, where CTCF blocks the promoter function of a neomycin resistance gene in the absence of ZBTB33. However, this is inconsistent to the results in this work, since mutation of the CTCF-binding site, which resulted in eliminating the possibility of CTCF to bind the *AREG* exon 2 sequence, led to a loss of promoter function. As a consequence, CTCF function in the promoter assays is not explained by its insulator function but rather by its function as a transcriptional activator as described by Vostrov *et al.*[114] To determine the role of CTCF on *AREG* expression, experiments could be done using siRNAs targeting CTCF. Chromatin-immunoprecipitation experiments should also be done to prove the CTCF- and ZBTB33-binding on the *AREG* exon 2. These experiments are currently in progress.

ZBTB33 as well as ZBTB4 were mainly identified as repressors of gene expression.[107,108] This also holds true for the *AREG* gene expression, since a reverse correlation was observed between the expression of both transcription factors and *AREG*. In literature, the function of ZBTB33 and ZBTB4 is mostly explained by their interaction with chromatin remodeling complexes.[108] Therefore, it is not clear if the effects seen in the promoter experiments are influenced by ZBTB33 or ZBTB4 at all, because the reporter plasmids are chromatin-free.

A hypothesis explaining the chromatin-free results might be as follows: An increased promoter function in the reverse-oriented *AREG* exon 2 sequence coupled to the Tata-binding motif or uncoupled, suggested a promoter function in reversed-orientation within the *AREG* gene. This might lead to an antisense transcript. If the antisense transcript interferes with the *AREG* gene expression, an inhibition of *AREG* gene expression would occur, independent of the chromatin structure. Therefore, it was further tested, if an antisense transcript exists, or not (see next section).

Figure 63:

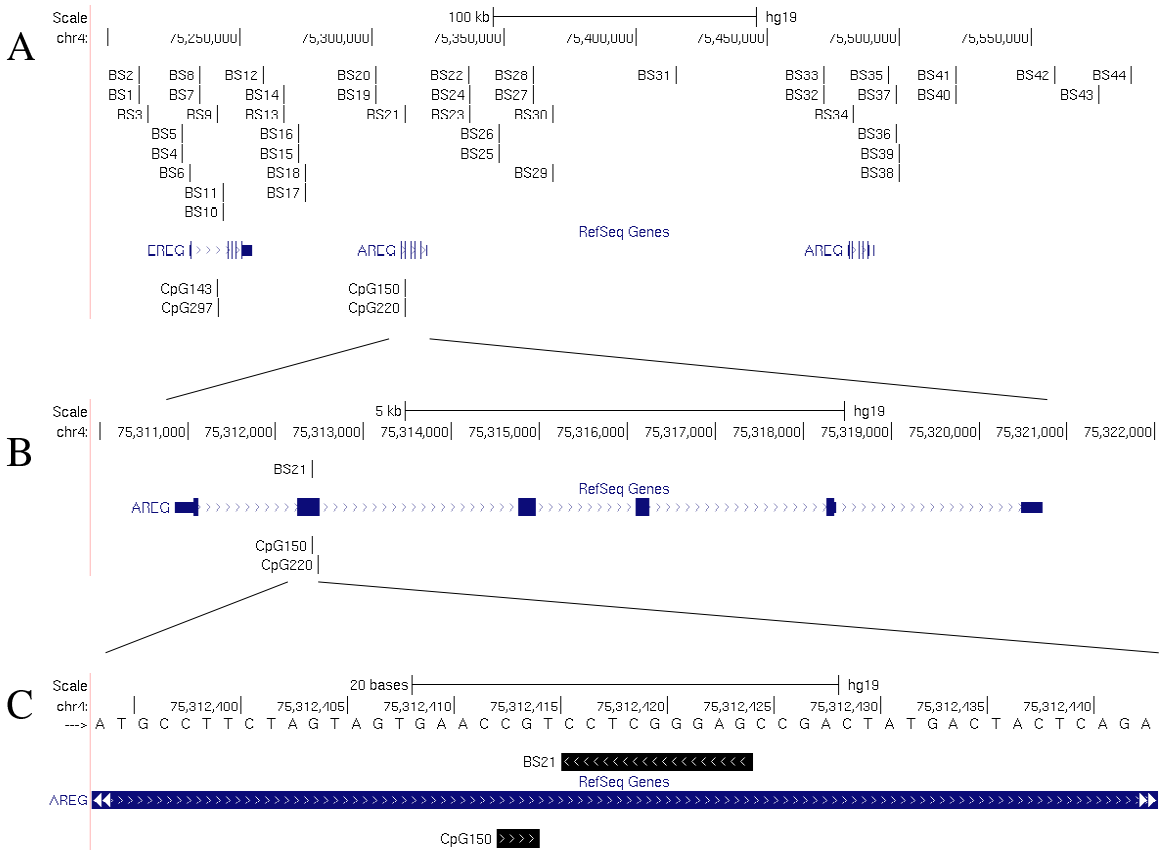


Figure 63: ZBTB33-binding sites identified at the *AREG* and *EREG* gene-locus: A: 44 unique ZBTB33-binding sites were identified at the *AREG* and *EREG* gene-locus, B: One ZBTB33-binding site is located at the *AREG* gene, C: This site is located closely to *AREG* CpG p150.

4.4 The ENCODE-dataset suggests an antisense transcript within the *AREG* gene

As shown previously, the *AREG* exon 2 sequence increased promoter function of the Tata-binding motif when coupled in reverse orientation to it, but not in forward orientation. This indicates that the *AREG* exon 2 could exert a regulatory function for transcription of an antisense transcript within the *AREG* gene. This hypothetical antisense RNA might be an enhancer-templated noncoding RNA as described in a review by Natoli and Andrau. The authors grouped these RNAs into 2 groups, 2d-eRNAs and 1d-eRNAs. Interestingly, the 1d-eRNAs, are defined to be transcribed unidirectional.[269]

A second hypothesis could be that the *AREG* exon 2 is a promoter for a long noncoding RNA (lncRNA), which interferes transcription of the *AREG* gene by being transcribed in reverse orientation. During transcription, the RNA-polymerase II traverses the *AREG* promoter which in consequence blocks the promoter.[270] One example

seen in mouse is described for the lncRNA *Airn*. By being transcribed in reverse orientation through another promoter, *Airn* can block the transcription of the protein-coding gene *Igf2r*.[271]

To address this hypothetical RNA, Northern blot experiments were performed (see section 3.8). The probe "E2-sense" was used as positive control to detect the *AREG* mRNA. Distinct signals being also as strong as expected, validated the experimental procedure and the probes used, because the antisense-specific probes were designed in the same way as the "E2-sense" probe. However, no distinct antisense product was identified using the probe "I1-antisense" and "E2-antisense. There was rather a strong signal spreading over the whole blot. Unspecific binding of the probes might be responsible for this result, because the lanes loaded with the marker showed similar strong signals as the lanes loaded with the samples. Also degraded RNA might lead to a signal spreading. Thus, Northern blot analyses could not identify the antisense RNA. Therefore, a second approach was performed to unravel a potential antisense transcript within the *AREG* gene.

CDNA was generated from cells with low or high *AREG* expression. Bisulfite-treatment of these cDNA enabled the detection of an antisense transcript within the *AREG* gene via a strand-specific PCR. Amplification products were detected, demonstrating the antisense sequence of *AREG* intron 1 and *AREG* exon 2 (see section 3.8.2). After sequencing, an unspecific PCR product could be ruled out. Several controls showed that the RNA used for cDNA synthesis was free of genomic DNA. Thus, the amplification product had to be derived from cDNA. Unfortunately, the assigned negative control (*UBE2D2*) was also positive for an antisense construct, although this was not expected and not indicated in the UCSC genome browser. Interestingly, an unspecific amplification product could be ruled out by sequencing, too. It remained controversial how the results should be interpreted. It might even possible, that an unknown aspect of the method itself might lead to false positive signals.

However, data acquired by the ENCODE-project[110] indicate also an antisense transcript starting at *AREG* exon 2 (see figure 64). The ENCODE project identified genome-wide histone modification patterns for different cell lines. The H3K27-acetylation e.g. is a mark for active regulatory elements, distinguishing active enhancers and promoters from inactive counterparts. In the *AREG* gene this mark is of course increased downstream the *AREG* promoter, but more interestingly, also upstream of exon 2 (see figure 64 A). This observation is further supported by the DNaseI hypersensitivity clusters, which show not only a strong peak at the promoter but also a weak peak at the exon 2 (see figure 64 A). Next, several transcription factor binding sites within the *AREG* exon 2 were identified by Chip-sequencing in the ENCODE project including chromodomain helicase DNA binding protein 2 (CHD2), the DNA endonuclease Rad2, the Tata-binding protein (TBP) and Z274 (see figure 64 C).[110] Interestingly, the RNA polymerase II (Pol2) was also found binding up- and downstream *AREG* exon 2 within two different cell lines

in independent data sets. Finally, H3K4-methylation, characteristic for open chromatin, showed an increased signal at the *AREG* exon 2 region for example in the K562 cell line (see figure 64 D). Therefore, most of the ENCODE-dataset suggest a promoter function at *AREG* exon 2.

Figure 64:

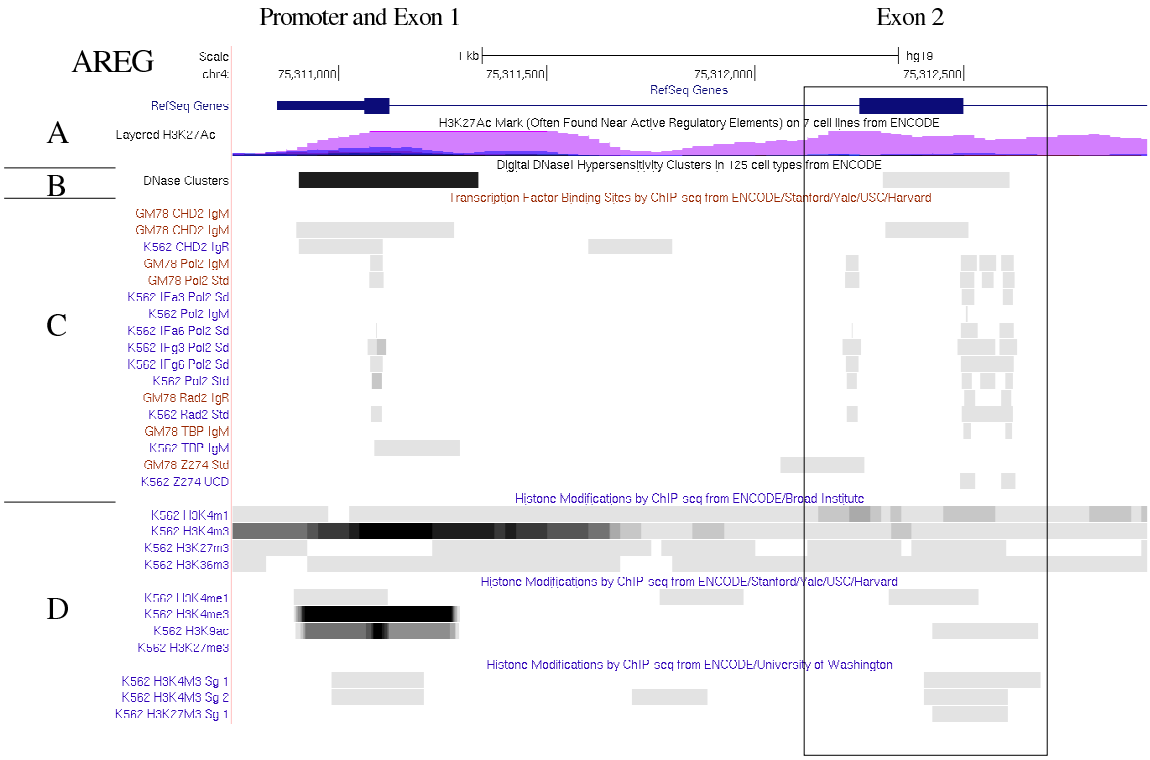


Figure 64: Selected ENCODE-data within the *AREG* gene: ENCODE-data were aligned to the *AREG* gene including exon 2. The UCSC genome browser was used for visualization. A: H3K27Ac-mark merged from 7 different cell lines; B: DNaseI hypersensitivity clusters merged from 125 cell lines; C: Selected transcription factor binding sites, identified by ChIP-seq; D: Histone methylation marks in K562 identified by three independent institutions within the ENCODE consortium. The black bordered area represents the *AREG* exon 2 region.

4.5 *AREG* and *EREG* expression and methylation of *AREG* CpG p150 and CpG p220 might be predictive markers for the outcome of EGFR-targeted therapies in *KRAS*-wildtype cells

In standard clinical settings, the *KRAS* mutation status of colorectal cancers is a selective negative marker for treatment with the EGFR inhibitor Cetuximab. It has been shown in several studies that only *KRAS*-wildtype patients can benefit from a therapy, since inhibition of EGFR might not inhibit downstream pathways when *KRAS* is mutated. In this work, Cetuximab was substituted by Erlotinib and Gefitinib in the *in vitro* experiments, because none of the cell lines showed any sensitivity towards Cetuximab. The reason was never resolved, but might include the microenvironment of the tumor. An increased immune-response upon Cetuximab treatment was for example detected *in vivo*,[272,273] which can never occur *in vitro*, because an immune system is not present in cell culture. Beside possible effects on the immune-system, Cetuximab also influences tumor growth by inhibiting the EGFR. It blocks the EGFR at the extra-cellular domain preventing EGFR-ligand induced dimerization and autophoshorylation of the receptor. This leads to a reduced activation of the downstream signaling-pathways. Erlotinib and Gefitinib inhibit the EGFR-function, too, but here an inhibition of the intra-cellular tyrosine kinase domains takes place preventing also autophosphorylation of the receptor. This also leads to a reduced activation of the downstream signaling-pathways. The results obtained by using these compounds should therefore be similar to the results obtained by Cetuximab treatment *in vivo*. Nevertheless, when interpreting the data, one should keep in mind that in *in vivo* settings, several other effects might also occur upon EGFR-targeted therapies.

Table 22: Comparison between AREG expression, intra-gene methylation and sensitivity towards Erlotinib and Gefitinib in five colorectal cell lines: The ΔCt-values of the *AREG* mRNA expression (see figure 3), and the methylation indices of CpG p150 and CpG p220 (MI, see figure 14) were averaged from four different timepoints. The sensitivity towards Erlotinib and Gefitinib was estimated from XTT-experiments (see section 3.9.1). -: cell line is resistant, +: cell line shows little response, ++: cell line shows intermediate response, +++: cell line shows strong response.

Cell line	AREG expression (mean ΔCt)	AREG CpG p150 / p220 (mean MI)	Sensitivity towards Erlotinib / Gefitinib
SW480	-3,87	0,23 / 0,32	++ / +++
RKO	-3,78	0,78 / 0,85	- / -
LIM1215	-1,42	0,57 / 0,64	+ / +
HCT116	0,20	0 / 0,11	+ / +
CaCO2	1,40	0 / 0,03	+++ / +++

Among the cell lines tested, RKO cells were completely resistant, LIM1215 and

HCT116 cells showed little response and SW480 as well as CaCO2 cells were highly sensitive towards Erlotinib or Gefitinib treatment (see section 3.9.1). Interestingly, HCT116 as well as SW480 cells have a *KRAS* mutation and therefore should not react upon treatment (see table 12). Since EGFR also influences other pathways directly, like the PI3K-pathway, inhibition of EGFR might also affect this pathway and might lead to a reduction of proliferation. Another example: RKO cells should be sensitive upon treatment since this cell line is *KRAS*-wildtype. However, this cell line has got a *BRAF* mutation. The BRAF protein acts downstream KRAS and also activates the MAPK-pathway independent of the EGFR-signal. Therefore, *BRAF* might be another candidate gene to predict the outcome of EGFR-targeted therapies and to select patients more stringent for treatment. A restrospective study using 113 patients showed that wildtype *BRAF* is necessary for successful Cetuximab or Panitumumab therapies.[274] However, due to its low abundance among all colorectal cancer patients leading to insignificant results in clinical studies, the predictive power of mutated *BRAF* could not be proven until now.

AREG and *EREG* expression was already described to correlate with the outcome of EGFR-targeted therapies in *KRAS*-wildtype patients,[194,195] or independent to the *KRAS*-status.[193] Nevertheless, in our cell lines, no correlation was observed between the endogenous *AREG* or *EREG* expression and the sensitivity towards Erlotinib and Gefitinib. HCT116 cells, LIM1215 and RKO cells were resistant or showed only little response, but in these cells *AREG* and *EREG* expression differed strongly (compare figures 3 and 41 A, 42 A and 43 A, see also table 22). Also, the sensitive cell lines, CaCO2 and SW480, showed very different *AREG* and *EREG* expression. However, in this study, only CaCO2 was *KRAS*-wildtype (see table 12) and had no other known mutations within the MAPK-pathway. This cell line was indeed sensitive to Erlotinib and Gefitinib and expressed the highest levels of *AREG* and *EREG* among the cells tested. Therefore, CaCO2, showed a correlation between *AREG* expression and sensitivity towards EGFR inhibitors. To increase the validity of this result, a future experiment would be to test several other *KRAS*-wildtype cell lines and compare the results and to modulate *AREG* expression in CaCO2 cells followed by sensitivity testing.

On the one hand, *AREG* expression correlated with the methylation of *AREG* CpG p150 and CpG p220 (see section 3.4). Cells with an MI of 0 for *AREG* CpG p150 and an MI of less than 0.2 for *AREG* CpG p220 belong to the high *AREG* expressing cells. Cells showing an MI greater than 0.1 for *AREG* CpG p150 and an MI greater than 0.2 for *AREG* CpG p220 belong to the *AREG* low and medium expressing cells. On the other hand, *AREG* expression did mainly not correlate with the sensitivity towards Erlotinib and Gefitinib. Therefore, it is unlikely that the methylation of *AREG* CpG p150 and CpG p220 correlated with the sensitivity towards the inhibitors. Interestingly, by taking a closer look to the methylation indices (MI), it was noticed that the resistant cell lines RKO and the low sensitive cell line LIM1215 showed the strongest methylation at the *AREG* CpGs (see table 22, MI >0.5). On the contrary, the most sensitive cell

line CaCO2 showed the lowest methylation index at the *AREG* CpG p150 and CpG p220. SW480 cells, showing an intermediate response towards Erlotinib treatment, were stronger methylated than CaCO2 cells but less methylated than RKO and LIM1215 cells at the *AREG* CpGs p150 and p220. Thus, methylation of the the *AREG* CpGs p150 and p220 correlated with the sensitivity towards Erlotinib with the exception of HCT116 cells and towards Gefitinib with the exception of HCT116 and SW480 cells. (see table 22). To summarize, while the *AREG* expression in cell lines is a poor predictive marker for the sensitivity towards EGFR inhibitors *in vitro*, methylation of *AREG* CpG p150 and CpG p220 showed a better correlation.

In the ColoNET consortium, patient-derived xenografts were tested, to compare the results obtained *in vitro* with *in vivo* data. A summary of the data, collected by Maria Rivera-Markelova and Alexandra Schwan, is given in the supplemental material 8.11. Here, in nine xenografts deriving from *KRAS*-wildtype colorectal cancers or metastases five were sensitive to Cetuximab. These five xenografts expressed the highest *AREG* mRNA and protein levels among all samples. Only one additional xenograft, which also expressed high levels of *AREG* was resistant. Eight *KRAS*-mutant xenografts were tested, too. Only two showed some response towards Cetuximab, but these two xenografts did not represent high *AREG* mRNA and protein expressing samples. Therefore, the results indicated that the *KRAS*-wildtype status, together with *AREG* expression, might indeed be predictive for EGFR-targeted therapies *in vivo*. Interestingly, none of the xenografts showed any methylation in the *AREG* CpGs p150 and p220 (data not shown). The reason for that could not be uncovered by subsequent experiments (data not shown). Therefore, methylation of *AREG* CpG p150 and CpG p220 could not be used to predict the outcome of EGFR-targeted therapies in this xenograft experiments.

4.6 Epigenetic treatment influences the outcome of EGFR-targeted therapies

HDACis were reported to influence the response towards EGFR-inhibition in several cancer models.[225,226,227,228,229] To test, if these findings might be adapted to our settings the cell lines SW480, RKO, LIM1215 and CaCO2 were treated with HDACis and afterwards treated with Erlotinib or Gefitinib. LIM1215 cells showed the strongest effect on sensitivity towards Erlotinib and Gefitinib after Valproat treatment (see figures 47 and 51). Since *AREG* and *EREG* expression increased upon HDACi-treatment, a consequence was to test, if AREG or EREG are responsible for the increased sensitivity. LIM1215 cells were transfected transiently or by lentiviral infection with an *AREG*-sequence containing plasmid. As shown in section 3.10, in none of the experiments, an increased sensitivity towards EGFR-inhibition was achieved although *AREG* protein amounts and mRNA-expression increased in a similar way than after HDACi-treatment (compare figures 6, 7 and 10 with 61). Direct application of recombinant AREG into the

supernatant did also not lead to an increased sensitivity towards Erlotinib or Gefitinib (data not shown). Although LIM1215 cells were described to be a *KRAS*-wildtype cell line,[275] in our laboratory a mutation within the *KRAS* gene was detected in exon 4 (see table 12). It might be possible that *AREG* over-expression alone will not affect sensitivity due to this *KRAS* mutation. It could also be possible that the effect of the HDACis onto sensitivity is independent of *AREG*. To clarify this problem, a low *AREG* expressing cell line would be necessary without any mutations within the MAPK-pathway components. However, such a cell line was not available during this work. Therefore, the direct contribution of AREG to the sensitivity towards EGFR inhibitors could not be proven by the experiments presented here. Nevertheless, in a publication by Ferraros *et al.*, the authors verified a causal relationship between *AREG* and *EREG* expression and response to Cetuximab in *KRAS*-wildtype cells.[276] In contrast to the experiments presented here, they studied the effects of *AREG* and *EREG* expression on sensitivity towards Cetuximab by decreasing the levels of these proteins. They showed that targeting either *AREG* or *EREG* gene expression with shRNAs led to decreased response rates towards Cetuximab.

A further interesting issue is that in LIM1215 cells HDACi-treatment increases sensitivity towards EGFR-targeted therapeutics *KRAS* mutation independently. A consequence of this experiment could be that HDACis might have the potential to overcome *KRAS* mutation dependent resistance towards EGFR-targeting therapeutics. But, if this is true, it is unclear why the HDACi-treatment did not affect the other *KRAS*-mutated, or even the *KRAS*-wildtype cell lines, in a similar way. One reason might be that the cell lines tested have different response rates towards the HDACis used. Also, the target HDACs might account for the dissimilar effectiveness of the HDACis. In LIM1215 cells, Cambinol, for example, did not increase sensitivity towards EGFR-inhibition, although Trichostatin A and Valproat did (see figure 47). Nevertheless, while TSA and Valproat are mainly targeting HDACs class I and II, Cambinol targets the HDAC Sirtuin I, a HDAC class III.[277] SAHA, in contrast, was not tolerated by LIM1215 at all. Additionally, it was tested if these findings can be applied *in vivo*. LIM1215 cells were injected in mice to create xenografts. But, in contrast to the *in vitro* results, HDACi treatment did not lead to an increase of senstitivity towards EGFR inhibitors (see figure 54).

As a second approach the mice, injected with LIM1215 cells, were treated with the DNA methyltransferase inhibitor 5-Azacytidine. By using the DNA methyltransferase inhibitor DAC in *in vitro* experiments, *AREG* and *EREG* expression increased (see figure 4). But, as already described, post-treatment procedures were not possible, since the cells died after replating. Nevertheless, in mice 5'Azacytidine treatment, which is also a DNA methyltransferase inhibitor, led to an increased response to Erlotinib (see figure 54). Therefore, in *in vivo* systems global epigenetic changes might also play a role for EGFR-treatment response. But, when testing the *AREG* and *EREG* expression in the xenografted cells, it was observed that treatment did not influence mRNA-expression of

AREG or *EREG* (see figure 55). Additionally, IHC-experiments showed that the AREG protein expression varied strongly among the xenografted cells treatment independently (see table 21 and figure 56). As a consequence *AREG* or *EREG* mRNA-expression as well as AREG protein amounts might not correlate with the increase of sensitivity in this experimental model or cannot be measured precisely. It was, for example, hardly possible to analyze the 5-Azacytidine treated samples, because the tumor growth was reduced dramatically upon treatment. Nevertheless, a decreased methylation index was observed in the *AREG* intragenic CpGs p150 and p220 as well as in the *EREG* intragenic CpG p297 upon 5-Azacytidine treatment alone or in combination with Erlotinib (see figure 57). Therefore, the epigenetic regions might function as predictive markers in combination therapies, when the expression of AREG cannot be evaluated. The lack between methylation and expression in the xenograft model has to be further addressed. Overall, the results indicate for clinical settings that patients, which do not respond to EGFR-targeted therapies might become sensitive after treatment with epigenetically interfering compounds by different, yet largely unknown mechanisms.

4.7 *AREG* CpG p150 and CpG p220 are differentially methylated in human tumor samples

Subsequent experiments were performed in the ColoNET-consortium to evaluate if the results of this work could also be applied to human tissues. As already described, *KRAS*-wildtype and mutant tumors were grown on mice and the xenografts were measured for their *AREG* mRNA and protein expression and for their response towards Cetuximab (see supplementary material 8.11). Similar to the cell lines, *AREG* was differentially expressed in human carcinoma-derived xenografts. Additionally the xenografts also responded differently to Cetuximab according to their mutation status and *AREG* expression (see supplemetary tables 27 and 28).

As a next approach, methylation of the identified intragenic CpGs in the *AREG* gene was measured in 5 different human tissue samples. Each tissue was separated by Florian Rossner (Charité Berlin, Institute for Pathology) into 4 parts using microdissection: normal epithelium, normal stroma, tumor epithelium and tumor stroma. The results are shown in supplementary material 8.12. Three of the 4 normal epithelium samples (1 sample failed) and 3 of 5 normal stroma samples showed high methylation within the tested CpGs (MI >0.5). Four of 5 tumor stroma samples showed also high methylation. Interestingly only 1 of 5 tumor epithelium samples showed high methylation for one CpG (CpG p220), which means that the tumor epithelium samples were mainly unmethylated at the tested CpGs, whereas in the other compartments the CpGs were mainly methylated. To verify the results, the approach was repeated with 20 additional human tissues. Additionally to the *AREG* CpGs p150 and p220 also the *EREG* CpGs p143 and p297 were tested. The boxplots of all methylation indices observed in

this experiment are shown in supplementary material 8.13. Similar to the results before, the methylation at the *AREG* CpGs were in general lower in the tumor epithelium compartments compared to the tumor stroma, the normal stroma or the normal epithelium. This was also seen for the methylation of the *EREG* CpGs. However, looking in detail to the samples, there are also 5 samples, which show a high methylation index greater than 0.6 at the *AREG* CpG p150 and CpG p220 in the tumor epithelium. In contrast, for *EREG* the highest methylation index is only 0.6. Nevertheless, it needs to be clarified to what extent the differential methylation of the *AREG* and *EREG* CpGs within the tumor epithelium samples and between the different tissue types are influenced by remaining tumor stroma cells within the tumor epithelium samples. It would also be interesting to clarify with corresponding patient data, if the methylation of *AREG* CpG p150 and CpG p220 within the tumor epithelium might also predict the outcome of EGFR-targeted therapies, as it was seen for the *AREG* expression in *in vivo* experiments using xenografts (see section 4.5). However, to test that, the number of tested samples has to be increased dramatically to ensure significant results.

4.8 Remarks on experimental designs

4.8.1 AREG and EREG were analyzed in this work

In this work *AREG* and *EREG* were chosen from all known EGFR-ligand genes, because they are commonly over-expressed in colorectal cancer cells.[56,61] They were also found to be potential predictive markers for EGFR-targeted therapies.[193,194,195] Together with *BTC*, *AREG* and *EREG* were discovered to be regulated via DNA methylation in an experiment which showed different mRNA-expression patterns in HCT116 cells and HCT116 derivatives, lacking DNMT1 and 3b.[240] All other EGFR-ligands were not identified to be regulated epigenetically.[240] Also, intragenic CpG-positions were identified for all three genes, which have different methylation patterns in different cell lines. However, in first tests using two cell lines, only the CpG-positions identified for the *AREG* and *EREG* genes were found to correlate with the expression of the genes. The identified CpG-positions within the *BTC* gene showed similar methylation although *BTC* expression differed in these cells (see supplemtary material 8.10). Therefore, *BTC* was omitted from analysis within this study. Nevertheless, when finding an intragenic CpG, whose methylation correlates with *BTC* expression, it would be interesting to consider BTC in future experiments.

4.8.2 Solvents influence experimental outcome

When comparing the solvent treated samples (DMSO, Ethanol) with untreated samples an interesting effect was seen in the proliferation experiments. In some experiments, cells grew faster or were more resistant to subsequent Erlotinib or Gefitinib treatment (compare e.g. figure 47 A and B). This effect was not consistent between different

experiments. The reason is unknown. However, one of the solvents, DMSO, used as control for Erlotinib and Gefitinib treatment, was reported to influence differentiation in erythroleukemia cells[278] by influencing the cell cycle,[279] enzyme activity[280] and DNA.[281] DMSO was also the starting point for the development of HDACis like SAHA.[213] Solvent controls were necessary to evaluate the effects of the inhibitors to the cells. Therefore, in each growth experiment, the effects of the inhibitors on proliferation were compared to the solvent treated cells since the same amount of solvent was present in each inhibitor approach. The solvent effect was then taken into consideration when interpreting the results.

4.8.3 Differences between *in vitro* and *in vivo* results

AREG and *EREG* mRNA-expression were increased upon HDACi- and DAC-treatment *in vitro*. In mice, no increase of *AREG* or *EREG* mRNA-expression was observable after treatment (see figure 55). In addition, all ΔCt-values, obtained by real-time PCR expression analysis of the LIM1215 cell-derived xenografts, were similar to the ΔCt-values of untreated LIM1215 cells *in vitro* (compare figures 3 and 55). Mouse derived RNA within the sample, which might lead to this result, could be ruled out by expression analysis of mouse-specific *AREG* and *EREG*. Neither of them was detected (data not shown), illustrating that the *AREG* mRNA measured within the xenografts derived completely from the human cells. The reason, why no increase was observed upon treatment could depend on the ratio between the human and mouse cells within the sample. However, the human-specific control gene expression should adjust the results to the input cell or RNA amount. Immunohistochemistry of the xenografts with an antibody targeting AREG revealed that also the protein levels of AREG differed between different samples of the same treatment group (see figure 56 and table 21). Together these results indicate that a strong heterogeneity of *AREG* expression in LIM1215 cells grown *in vivo* render a quantitative evaluation of the effects of the epigenetic compounds onto sensitivity towards EGFR inhibitors impossible. Furthermore, contributing effects of mouse stromal cells onto *AREG* expression could not be ruled out.

CpG-methylation analysis was performed by Dr. Sascha Tierling (University of Saarbrücken), using the methylation-dependent single-nucleotide primer extension approach (msSNuPE). The methylation of *AREG* CpG p150 and CpG p220 was different between the LIM1215 cells *in vitro* and the LIM1215 derived xenografts (compare figures 14 and 57). *In vitro*, the MI ranged between 0.4 in CpG p150 and 0.6 in CpG p220, whereas in the xenografts the MI ranged from 0.2 in CpG p150 to 0.4 in CpG p220. Interestingly the MIs of the *EREG* CpGs p143 and p297 did not differ strongly between LIM1215 cells *in vitro* and *in vivo* (compare figures 14 and 57).

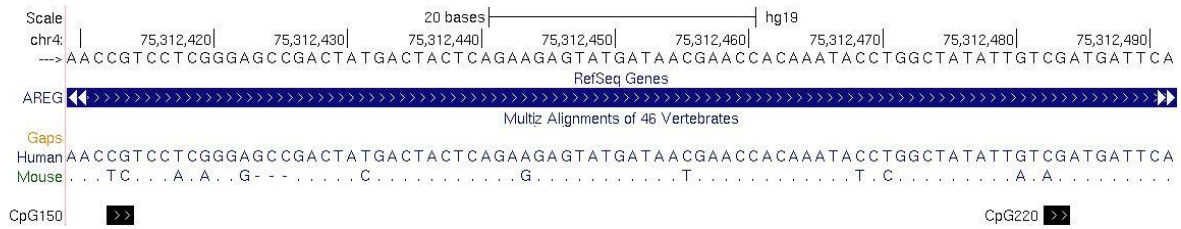
To address this problem, the *AREG* gene sequence of the human genome was compared to the sequence of the mouse genome. In figure 65 is shown that the sequences of *AREG* exon 2 are very similar in human and mouse. However, the *AREG* CpG

p150 is changed in the mouse genome to a TC-dinucleotide. After bisulfite conversion, unmethylated human cells and all mouse cells have binding sites within their DNA for the msSNuPE-primer detecting the unmethylated state of the *AREG* CpG p150. The reason is that bisulfite-treatment necessary for the msSNuPE approach will convert the Cytosine in the unmethylated human *AREG* CpG p150 to an Uracil, which is a derivative of Thymine. As a consequence every remnant mouse cell within the xenografted sample will decrease the MI-value of *AREG* CpG p150. Additionally, there is an Adenine in the mouse gene at the position of the cytosine of *AREG* CpG p220, which might also influence the MI.

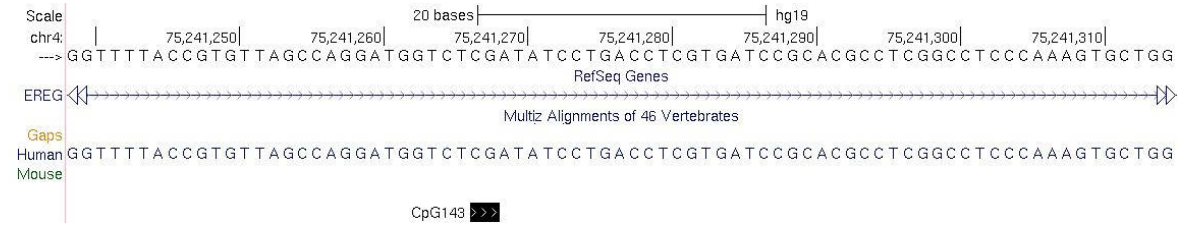
In contrast to *AREG*, when comparing the *EREG* gene sequences between human and mouse, it was observed that there is a strong difference within the region of *EREG* CpG p143 and p297 (see figure 65). In the region of *EREG* CpG p143, there is no consensus sequence within the mouse gene, in the region of *EREG* CpG p297, the sequence is very different between human and mouse. As a consequence, the msSNuPE-primers, which are used for the detection of these CpGs, only bind to human DNA. Therefore, the results between *in vitro* and *in vivo* should be similar as it was seen in the experiments.

Figure 65:

AREG CpGs p150 and p220:



EREG CpG p143:



EREG CpG p297:

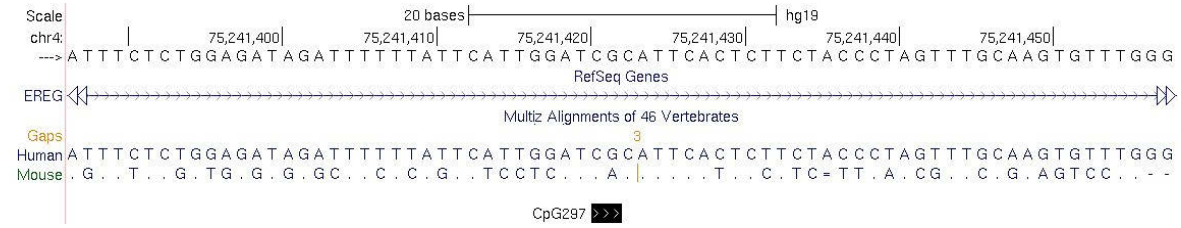


Figure 65: Sequence similarities at the *AREG* and *EREG* genes between human and mouse: The human genome hg19 (human) and the mouse genome mm9 (mouse) were compared at the genloci within the *AREG* and *EREG* genes where the intragenic CpGs of interest are located. (.): sequence is equal, (-): deletions within the mouse sequence, (N): base substitutions between mouse and human sequence, ( ): no sequence similarities.

4.9 Outlook

AREG gene expression is regulated by an epigenetic mechanism involving the intragenic CpGs within *AREG* exon 2. By comparing the results of promoter function analysis in different cell lines, bioinformatic approaches, and strand-specific PCR a hypothesis was made, how *AREG* exon 2 influences *AREG* expression (see figure 66). In this hypothesis, the methylation-dependent transcription factors CTCF, ZBTB33 and ZBTB4 play an important role. They interact methylation-dependently with the *AREG* exon 2 and suppress *AREG* expression. The gene suppression might be achieved either by interacting with histone modifying enzymes or by supporting the transcription of an antisense-RNA which interferes with the *AREG* transcription. Deregulating these transcription factors in the cell lines tested by over-expression, shRNA mediated down-regulation or by chemicals, followed by *AREG* expression analysis or promoter function analysis of *AREG* exon 2 would in consequence lead to a gain of information verifying or falsifying the importance

of ZBTB33 and CTCF on the *AREG* expression. Also chromatin-immunoprecipitation experiments need to be made to investigate if the transcription factors directly influence *AREG* expression by binding to the methylated or unmethylated *AREG* exon 2. Stronger efforts are also necessary to address the possible antisense-RNA by different methods as ribonuclease-protection-assays or other strand-discriminating methods. These experiments are currently in progress.

Xenograft experiments revealed that *AREG* high expressing *KRAS*-wildtype colorectal tumors show a higher senitivity to Cetuximab than *AREG* low expressing tumors (see supplementary material 8.11). As a consequence, high *AREG* expression or low methylation of *AREG* CpG p150 and p220 might be predictive markers in combination with wildtype *KRAS* to predict a positive outcome of EGFR-targeted therapies.

A next step would be to verify the correlation between *AREG* mRNA and protein expression and the methylation of *AREG* CpG p150 and CpGp220 *in vivo*. When using material with known *KRAS* mutation states and response rates towards EGFR-targeted therapies, it could further be tested, if there is a correlation between these characteristics and the methylation of the intragenic *AREG* CpGs. So, the predictive power of *AREG* expression or its intragene methylation could be evaluated.

Further experiments within this study showed that HDACi treatment *in vitro* and 5-Azacytidine treatment *in vivo* led to an increased sensitivity towards EGFR inhibitors, respectively. Interestingly, this effect occurred in the *KRAS* mutated cell line LIM1215. Additionally, these results were accompanied by increases in *AREG* expression or decreases in *AREG* exon 2 methylation. It would be interesting if this result might also be achieved in human tumor samples. It would be a milestone in cancer therapy, if non-responding patients might become sensitive to EGFR-targeting therapeutics *KRAS* mutation independently by pre-treating the patients with epigenetic compounds like HDACis. *AREG* expression or its intragenic methylation might then be a means of monitoring the process of becoming sensitive. First experiments to test the situation *in vivo* might be realized by xenografting tumor cells with different *KRAS* mutation states into mice and monitor their growth upon treatment with epigenetic compounds in combination with EGFR inhibitors. Efforts to realize these experiments are currently in progress, too.

Figure 66:

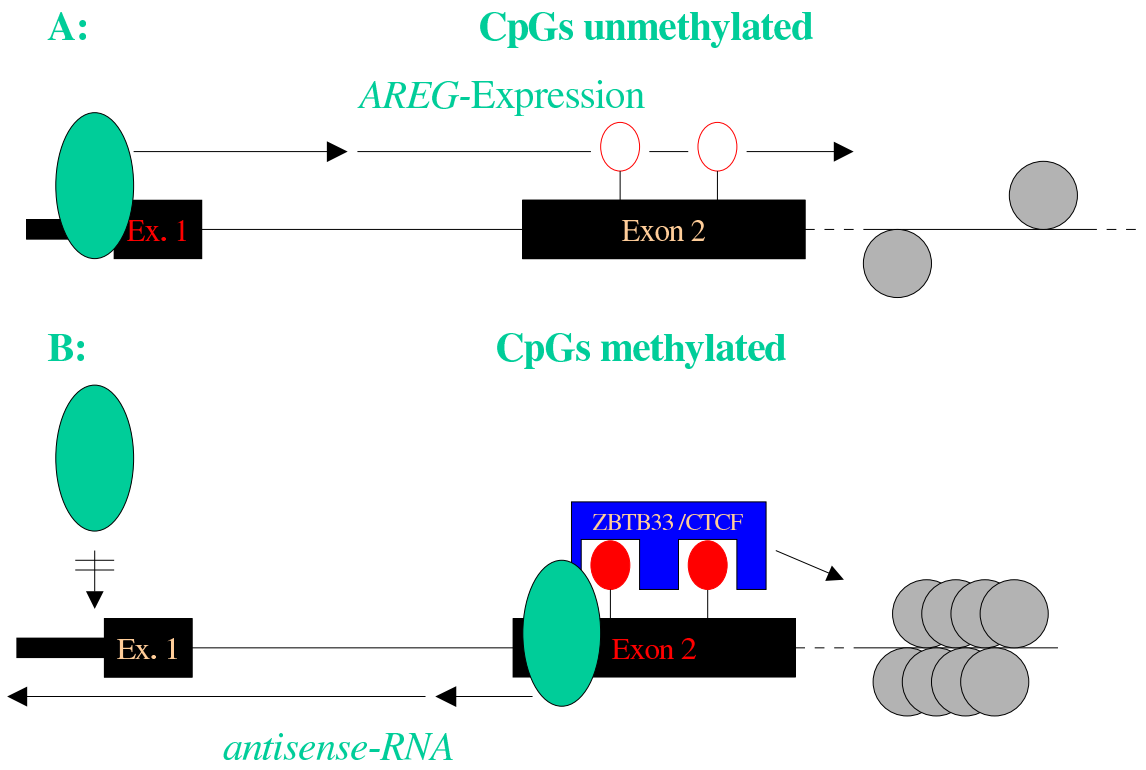


Figure 66: Hypothetic influence of *AREG* exon 2 on *AREG* expression: A: Normal transcription of the *AREG* gene, when the CpGs within *AREG* exon 2 are unmethylated. B: When the CpGs within *AREG* exon 2 are methylated, ZBTB33, CTCF or a complex containing both proteins is able to bind. As a consequence it closes the chromatin structure e.g. by interaction with chromatin remodeling complexes. ZBTB33 and CTCF also influence the RNAPII-dependent transcription of an antisense transcript. By traversing the *AREG* promoter, the antisense RNA anticipates binding of RNAPII to the *AREG* promoter. Empty red circles: unmethylated CpGs within *AREG* exon 2; full red circles: methylated CpGs. Black lines: DNA sequence; black bars: exons of the *AREG* gene, Green filled circle: RNAPII; blue form: CTCF and ZBTB33 interacting with *AREG* exon 2. Grey full circles: nucleosomes. Arrows indicate the orientation of transcription.

5 Summary

Amphiregulin and Epiregulin are ligands of the Epidermal Growth Factor Receptor (EGFR) whose expression correlates with a positive EGFR-targeted therapy response in colorectal cancer.[193,194,195] Aim of this work was to define the influence of epigenetic mechanisms on *AREG* and *EREG* gene expression. It could be shown that *AREG* and *EREG* are differentially expressed in a set of colorectal cancer cell lines and that the expression of both genes increases after treatment with epigenetically interfering compounds such as DNA methyltransferase inhibitors and histone deacetylase inhibitors. Therefore, the epigenetic mechanism regulating *AREG* and *EREG* gene expression was verified. Methylation analysis in five cell lines showed that the promoters of both genes were unmethylated, except in one cell line for the *EREG* promoter. Furthermore, short intragenic regions within both genes were identified to be differentially methylated. For *AREG*, this region is located within exon 2, indicating an uncommon epigenetic regulatory mechanism. Promoter function analyses showed that the region containing *AREG* exon 2 harbor methylation- and orientation dependent promoter function and they suggested CTCF, an MDB-protein, to be involved in the promoter function. A reduction of promoter function of reporter plasmids was observed upon mutation of a CTCF-binding site within the *AREG* exon 2 sequence. Expression analysis experiments suggested also ZBTB33, another MDB-protein, to be involved in *AREG* regulation. *ZBTB33* was differentially expressed in the cells and the expression correlated inversely with the *AREG* expression. Further experiments modulating the *ZBTB33* expression as well as a bioinformatic analyses identifying a ZBTB33 binding site within *AREG* exon 2 strengthen the hypothesis that also ZBTB33 is involved in *AREG* regulation.
It was also shown in this work that LIM1215 cells treated with HDACis were more sensitive towards EGFR inhibitors *in vitro*. This effect was accompanied by an increased *AREG* and *EREG* expression. *In vivo*, an increased sensitivity towards EGFR inhibitors was achieved in LIM1215 cells by treatment with a DNA methyltransferase inhibitor. Here the effect was accompanied by a reduced methylation within the *AREG* and *EREG* intragenic CpGs. Together, the results suggested a new possibility to potentially make EGFR-targeted therapy resistant patients suitable for EGFR-targeted therapies by epigenetic compound treatment. In that case *AREG* as well as *EREG* might be predictive markers to evaluate the effect of the epigenetic compounds during therapy.

6 Zusammenfassung

Amphiregulin und Epiregulin sind Liganden des epidermalen Wachstumsfaktor Rezeptors (EGFR), deren Expression mit einem positiven EGFR-zielgerichtetem Therapieansprechen in Darmkrebs korreliert.[193,194,195] Ziel dieser Arbeit war es, einen epigenetischen Einfluss auf die *AREG* und *EREG* Expression zu klären. Es wurde gezeigt, dass *AREG* und *EREG* in verschiedenen kolorektalen Krebszelllinien differenziell exprimiert sind, und dass die Expression beider Gene durch Substanzen, die mit epigenetischen Prozessen interferieren, zum Beispiel DNA-Methyltransferaseinhibitoren oder Histondeazetylaseinhibitoren, erhöht werden kann. Dadurch konnte der epigenetische Einfluss auf die *AREG* und *EREG* Expression bestätigt werden. Eine Analyse der DNA-Methylierung in fünf Zelllinien zeigte, dass die Promotoren beider Gene, bis auf eine Ausnahme in *EREG*, unmethyliert vorlagen. Hingegen wurden kurze Regionen innerhalb der kodierenden Bereiche der Gene als differentiell methyliert identifiziert. Im *AREG* Gen liegt diese Region im Exon 2, was auf einen ungewöhnlichen epigenetisch gesteuerten Regulationsmechanismus hindeutet. Promotorfunktionsanalysen zeigten, dass die *AREG* Exon 2 Region eine methylierungsabhängige und orientierungsabhängige Promoterfunktion hat, in die das MDB-Protein CTCF involviert sein könnte. Die Mutation einer CTCF-Bindungsstelle in der *AREG* Exon 2 Sequenz einiger Reporterplasmide führte zur Veringerung der Promotorfunktion dieser Plasmide. Expressionsanalysen wiesen jedoch darauf hin, dass auch ZBTB33, ein anderes MDB-Protein, in die *AREG* Regulation involviert sein könnte. Die *ZBTB33* Expression war in den getesteten Zelllinien differentiell und sie war negativ mit der *AREG* Expression korreliert. Weiterführende Experimente zur Modulation der *ZBTB33* Expression und eine bioinformatorische Untersuchung, die eine neue ZBTB33-Bindungsstelle im *AREG* Exon 2 identifizierte, stärkten die Hypothese, dass ZBTB33 eine wichtige Rolle in der *AREG* Regulation spielt.

Des Weiteren wurde gezeigt, dass die Behandlung der Zelllinie LIM1215 mit Histondeazetylaseinhibitoren im *in vitro* Experiment zu einer Erhöhung der Sensitivität gegenüber EGFR-zielgerichteten Medikamenten führt. Begleitet wurde der Effekt mit einer Erhöhung der *AREG*- und *EREG*-Expression. Im *in vivo* Versuch konnte die Sensitivität von LIM1215 Zellen gegenüber EGFR-zielgerichteten Medikamenten durch die Behandlung mit DNA-Methyltransferaseinhibitoren erhöht werden. Begleitet wurde der Effekt hier mit einer Verringerung der Methylierung der untersuchten *AREG*- und *EREG*-intragenischen CpGs. Diese Ergebnisse zeigen eine neue Möglichkeit auf, wie Patienten, die resistent gegenüber EGFR-zielgerichteten Therapien sind, möglicherweise sensitiv gemacht werden können. In diesem Fall könnten sowohl *AREG* als auch *EREG* als prädiktive Marker eingesetzt werden, um den Effekt der die Epigenetik beeinflussenden Substanzen zu evaluieren.

References

[1] Kaatsch, P., Spix, C., Katalinic, A. and Hentschel, S. (2012). Krebs in Deutschland 2007/2008, vol. 8. Robert Koch-Institut (Hrsg) und die Gesellschaft der epidemiologischen Krebsregister in Deutschland e.V. (Hrsg)., Berlin.

[2] Potter, J.D. (1999). Colorectal cancer: molecules and populations. *J Natl Cancer Inst* 91(11):916–932.

[3] Slattery, M.L. (2000). Diet, lifestyle, and colon cancer. *Semin Gastrointest Dis* 11(3):142–146.

[4] Huxley, R.R., Ansary-Moghaddam, A., Clifton, P., Czernichow, S., Parr, C.L. and Woodward, M. (2009). The impact of dietary and lifestyle risk factors on risk of colorectal cancer: a quantitative overview of the epidemiological evidence. *Int J Cancer* 125(1):171–180.

[5] Kinzler, K.W. and Vogelstein, B. (1996). Lessons from hereditary colorectal cancer. *Cell* 87(2):159–170.

[6] Rubinfeld, B., Souza, B., Albert, I., Müller, O., Chamberlain, S.H., Masiarz, F.R., Munemitsu, S. and Polakis, P. (1993). Association of the APC gene product with beta-catenin. *Science* 262(5140):1731–1734.

[7] Fearon, E.R. (2011). Molecular genetics of colorectal cancer. *Annu Rev Pathol* 6:479–507.

[8] Papadopoulos, N., Nicolaides, N.C., Wei, Y.F., Ruben, S.M., Carter, K.C., Rosen, C.A., Haseltine, W.A., Fleischmann, R.D., Fraser, C.M. and Adams, M.D. (1994). Mutation of a mutL homolog in hereditary colon cancer. *Science* 263(5153):1625–1629.

[9] Bronner, C.E., Baker, S.M., Morrison, P.T., Warren, G., Smith, L.G., Lescoe, M.K., Kane, M., Earabino, C., Lipford, J. and Lindblom, A. (1994). Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. *Nature* 368(6468):258–261.

[10] Fishel, R., Lescoe, M.K., Rao, M.R., Copeland, N.G., Jenkins, N.A., Garber, J., Kane, M. and Kolodner, R. (1993). The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell* 75(5):1027–1038.

[11] Leach, F.S., Nicolaides, N.C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomäki, P., Sistonen, P., Aaltonen, L.A. and Nyström-Lahti, M. (1993). Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell* 75(6):1215–1225.

[12] Grady, W.M. (2004). Genomic instability and colon cancer. *Cancer Metastasis Rev* 23(1-2):11–27.

[13] Wood, L.D., Parsons, D.W., Jones, S., Lin, J., Sjöblom, T., Leary, R.J., Shen, D., Boca, S.M., Barber, T., Ptak, J. et al. (2007). The genomic landscapes of human breast and colorectal cancers. *Science* 318(5853):1108–1113.

[14] Baker, S.J., Preisinger, A.C., Jessup, J.M., Paraskeva, C., Markowitz, S., Willson, J.K., Hamilton, S. and Vogelstein, B. (1990). p53 gene mutations occur in combination with 17p allelic deletions as late events in colorectal tumorigenesis. *Cancer Res* 50(23):7717–7722.

[15] Teodoro, J.G., Evans, S.K. and Green, M.R. (2007). Inhibition of tumor angiogenesis by p53: a new role for the guardian of the genome. *J Mol Med (Berl)* 85(11):1175–1186.

[16] Mendelsohn, J. and Baselga, J. (2000). The EGF receptor family as targets for cancer therapy. *Oncogene* 19(56):6550–6565.

[17] Boulougouris, P. and Elder, J. (2001). Epidermal growth factor receptor structure, regulation, mitogenic signalling and effects of activation. *Anticancer Res* 21(4A):2769–2775.

[18] Normanno, N., Bianco, C., Strizzi, L., Mancino, M., Maiello, M.R., Luca, A.D., Caponigro, F. and Salomon, D.S. (2005). The ErbB receptors and their ligands in cancer: an overview. *Curr Drug Targets* 6(3):243–257.

[19] Graus-Porta, D., Beerli, R.R., Daly, J.M. and Hynes, N.E. (1997). ErbB-2, the preferred heterodimerization partner of all ErbB receptors, is a mediator of lateral signaling. *EMBO J* 16(7):1647–1655.

[20] Baldus, S.E., Schaefer, K.L., Engers, R., Hartleb, D., Stoecklein, N.H. and Gabbert, H.E. (2010). Prevalence and heterogeneity of KRAS, BRAF, and PIK3CA mutations in primary colorectal adenocarcinomas and their corresponding metastases. *Clin Cancer Res* 16(3):790–799.

[21] Palomba, G., Colombino, M., Contu, A., Massidda, B., Baldino, G., Pazzola, A., Ionta, M., Capelli, F., Trova, V., Sedda, T. et al. (2012). Prevalence of KRAS, BRAF, and PIK3CA somatic mutations in patients with colorectal carcinoma may vary in the same population: clues from Sardinia. *J Transl Med* 10:178.

[22] O'dwyer, P.J. and Benson, A.B. (2002). Epidermal growth factor receptor-targeted therapy in colorectal cancer. *Semin Oncol* 29(5 Suppl 14):10–17.

[23] Busser, B., Sancey, L., Brambilla, E., Coll, J.L. and Hurbin, A. (2011). The multiple roles of amphiregulin in human cancer. *Biochim Biophys Acta* 1816(2):119–131.

[24] Shoyab, M., McDonald, V.L., Bradley, J.G. and Todaro, G.J. (1988). Amphiregulin: a bifunctional growth-modulating glycoprotein produced by the phorbol 12-myristate 13-acetate-treated human breast adenocarcinoma cell line MCF-7. *Proc Natl Acad Sci U S A* 85(17):6528–6532.

[25] Toyoda, H., Komurasaki, T., Uchida, D., Takayama, Y., Isobe, T., Okuyama, T. and Hanada, K. (1995). Epiregulin. A novel epidermal growth factor with mitogenic activity for rat primary hepatocytes. *J Biol Chem* 270(13):7495–7500.

[26] Brown, C.L., Meise, K.S., Plowman, G.D., Coffey, R.J. and Dempsey, P.J. (1998). Cell surface ectodomain cleavage of human amphiregulin precursor is sensitive to a metalloprotease inhibitor. Release of a predominant N-glycosylated 43-kDa soluble form. *J Biol Chem* 273(27):17258–17268.

[27] Sunnarborg, S.W., Hinkle, C.L., Stevenson, M., Russell, W.E., Raska, C.S., Peschon, J.J., Castner, B.J., Gerhart, M.J., Paxton, R.J., Black, R.A. et al. (2002). Tumor necrosis factor-alpha converting enzyme (TACE) regulates epidermal growth factor receptor ligand availability. *J Biol Chem* 277(15):12838–12845.

[28] Isokane, M., Hieda, M., Hirakawa, S., Shudou, M., Nakashiro, K., Hashimoto, K., Hamakawa, H. and Higashiyama, S. (2008). Plasma-membrane-anchored growth factor pro-amphiregulin binds A-type lamin and regulates global transcription. *J Cell Sci* 121(Pt 21):3608–3618.

[29] Sahin, U., Weskamp, G., Kelly, K., Zhou, H.M., Higashiyama, S., Peschon, J., Hartmann, D., Saftig, P. and Blobel, C.P. (2004). Distinct roles for ADAM10 and ADAM17 in ectodomain shedding of six EGFR ligands. *J Cell Biol* 164(5):769–779.

[30] Singh, A.B. and Harris, R.C. (2005). Autocrine, paracrine and juxtacrine signaling by EGFR ligands. *Cell Signal* 17(10):1183–1193.

[31] Tsark, E.C., Adamson, E.D., Withers, G.E. and Wiley, L.M. (1997). Expression and function of amphiregulin during murine preimplantation development. *Mol Reprod Dev* 47(3):271–283.

[32] Kariagina, A., Xie, J., Leipprandt, J.R. and Haslam, S.Z. (2010). Amphiregulin mediates estrogen, progesterone, and EGFR signaling in the normal rat mammary gland and in hormone-dependent rat mammary cancers. *Horm Cancer* 1(5):229–244.

[33] Ejskjaer, K., Srensen, B.S., Poulsen, S.S., Mogensen, O., Forman, A. and Nex, E. (2005). Expression of the epidermal growth factor system in human endometrium during the menstrual cycle. *Mol Hum Reprod* 11(8):543–551.

[34] Das, S.K., Chakraborty, I., Paria, B.C., Wang, X.N., Plowman, G. and Dey, S.K. (1995). Amphiregulin is an implantation-specific and progesterone-regulated gene in the mouse uterus. *Mol Endocrinol* 9(6):691–705.

[35] Park, J.Y., Su, Y.Q., Ariga, M., Law, E., Jin, S.L.C. and Conti, M. (2004). EGF-like growth factors as mediators of LH action in the ovulatory follicle. *Science* 303(5658):682–684.

[36] Tørring, N., Sørensen, B.S., Bösch, S.T., Klocker, H. and Nexø, E. (1998). Amphiregulin is expressed in primary cultures of prostate myofibroblasts, fibroblasts, epithelial cells, and in prostate tissue. *Prostate Cancer Prostatic Dis* 1(5):262–267.

[37] Sakurai, H., Tsukamoto, T., Kjelsberg, C.A., Cantley, L.G. and Nigam, S.K. (1997). EGF receptor ligands are a large fraction of in vitro branching morphogens secreted by embryonic kidney. *Am J Physiol* 273(3 Pt 2):F463–F472.

[38] Schuger, L., Johnson, G.R., Gilbride, K., Plowman, G.D. and Mandel, R. (1996). Amphiregulin in lung branching morphogenesis: interaction with heparan sulfate proteoglycan modulates cell proliferation. *Development* 122(6):1759–1767.

[39] Shiraishi, K. and Matsuyama, H. (2012). Local expression of epidermal growth factor-like growth factors in human testis and its role in spermatogenesis. *J Androl* 33(1):66–73.

[40] Qin, L., Tamasi, J., Raggatt, L., Li, X., Feyen, J.H.M., Lee, D.C., Dicicco-Bloom, E. and Partridge, N.C. (2005). Amphiregulin is a novel growth factor involved in normal bone development and in the cellular response to parathyroid hormone stimulation. *J Biol Chem* 280(5):3974–3981.

[41] Qin, L. and Partridge, N.C. (2005). Stimulation of amphiregulin expression in osteoblastic cells by parathyroid hormone requires the protein kinase A and cAMP response element-binding protein signaling pathway. *J Cell Biochem* 96(3):632–640.

[42] Falk, A. and Frisén, J. (2002). Amphiregulin is a mitogen for adult neural stem cells. *J Neurosci Res* 69(6):757–762.

[43] Nilsson, A., Moller, K., Dahlin, L., Lundborg, G. and Kanje, M. (2005). Early changes in gene expression in the dorsal root ganglia after transection of the sciatic nerve; effects of amphiregulin and PAI-1 on regeneration. *Brain Res Mol Brain Res* 136(1-2):65–74.

[44] Zhang, Y., Kobayashi, T., Hayashi, Y., Yoshioka, R., Shiraishi, A., Shirasawa, S., Higashiyama, S. and Ohashi, Y. (2012). Important role of epiregulin in inflammatory responses during corneal epithelial wound healing. *Invest Ophthalmol Vis Sci* 53(4):2414–2423.

[45] Kim, J.M., Bak, E.J., Chang, J.Y., Kim, S.T., Park, W.S., Yoo, Y.J. and Cha, J.H. (2011). Effects of HB-EGF and epiregulin on wound healing of gingival cells in vitro. *Oral Dis* 17(8):785–793.

[46] Takahashi, M., Hayashi, K., Yoshida, K., Ohkawa, Y., Komurasaki, T., Kitabatake, A., Ogawa, A., Nishida, W., Yano, M., Monden, M. et al. (2003). Epiregulin as a major autocrine/paracrine factor released from ERK- and p38MAPK-activated vascular smooth muscle cells. *Circulation* 108(20):2524–2529.

[47] LeJeune, S., Leek, R., Horak, E., Plowman, G., Greenall, M. and Harris, A.L. (1993). Amphiregulin, epidermal growth factor receptor, and estrogen receptor expression in human primary breast cancer. *Cancer Res* 53(15):3597–3602.

[48] Normanno, N., Kim, N., Wen, D., Smith, K., Harris, A.L., Plowman, G., Colletta, G., Ciardiello, F. and Salomon, D.S. (1995). Expression of messenger RNA for amphiregulin, heregulin, and cripto-1, three new members of the epidermal growth factor family, in human breast carcinomas. *Breast Cancer Res Treat* 35(3):293–297.

[49] Fontanini, G., Laurentiis, M.D., Vignati, S., Chin, S., Lucchi, M., Silvestri, V., Mussi, A., Placido, S.D., Tortora, G., Bianco, A.R. et al. (1998). Evaluation of epidermal growth factor-related growth factors and receptors and of neoangiogenesis in completely resected stage I-IIIA non-small-cell lung cancer: amphiregulin and microvessel count are independent prognostic indicators of survival. *Clin Cancer Res* 4(1):241–249.

[50] Hurbin, A., Dubrez, L., Coll, J.L. and Favrot, M.C. (2002). Inhibition of apoptosis by amphiregulin via an insulin-like growth factor-1 receptor-dependent pathway in non-small cell lung cancer cell lines. *J Biol Chem* 277(51):49127–49133.

[51] Castillo, J., Erroba, E., Perugorría, M.J., Santamaría, M., Lee, D.C., Prieto, J., Avila, M.A. and Berasain, C. (2006). Amphiregulin contributes to the transformed phenotype of human hepatocellular carcinoma cells. *Cancer Res* 66(12):6129–6138.

[52] Kuramochi, H., Nakajima, G., Kaneko, Y., Nakamura, A., Inoue, Y., Yamamoto, M. and Hayashi, K. (2012). Amphiregulin and Epiregulin mRNA expression in primary colorectal cancer and corresponding liver metastases. *BMC Cancer* 12:88.

[53] Bostwick, D.G., Qian, J. and Maihle, N.J. (2004). Amphiregulin expression in prostatic intraepithelial neoplasia and adenocarcinoma: a study of 93 cases. *Prostate* 58(2):164–168.

[54] Ebert, M., Yokoyama, M., Kobrin, M.S., Friess, H., Lopez, M.E., Büchler, M.W., Johnson, G.R. and Korc, M. (1994). Induction and expression of amphiregulin in human pancreatic cancer. *Cancer Res* 54(15):3959–3962.

[55] Oikawa, T., Hitomi, J., Kono, A., Kaneko, E. and Yamaguchi, K. (1995). Frequent expression of genes for receptor tyrosine kinases and their ligands in human pancreatic cancer cells. *Int J Pancreatol* 18(1):15–23.

[56] Saeki, T., Stromberg, K., Qi, C.F., Gullick, W.J., Tahara, E., Normanno, N., Ciardiello, F., Kenney, N., Johnson, G.R. and Salomon, D.S. (1992). Differential immunohistochemical detection of amphiregulin and cripto in human normal colon and colorectal tumors. *Cancer Res* 52(12):3467–3473.

[57] Ciardiello, F., Kim, N., Saeki, T., Dono, R., Persico, M.G., Plowman, G.D., Garrigues, J., Radke, S., Todaro, G.J. and Salomon, D.S. (1991). Differential expression of epidermal growth factor-related proteins in human colorectal tumors. *Proc Natl Acad Sci U S A* 88(17):7792–7796.

[58] Cook, P.W., Pittelkow, M.R., Keeble, W.W., Graves-Deal, R., Coffey, R.J. and Shipley, G.D. (1992). Amphiregulin messenger RNA is elevated in psoriatic epidermis and gastrointestinal carcinomas. *Cancer Res* 52(11):3224–3227.

[59] McIntyre, E., Blackburn, E., Brown, P.J., Johnson, C.G. and Gullick, W.J. (2010). The complete family of epidermal growth factor receptors and their ligands are co-ordinately expressed in breast cancer. *Breast Cancer Res Treat* 122(1):105–110.

[60] Sunaga, N., Kaira, K., Imai, H., Shimizu, K., Nakano, T., Shames, D.S., Girard, L., Soh, J., Sato, M., Iwasaki, Y. et al. (2012). Oncogenic KRAS-induced epiregulin overexpression contributes to aggressive phenotype and is a promising therapeutic target in non-small-cell lung cancer. *Oncogene* .

[61] Baba, I., Shirasawa, S., Iwamoto, R., Okumura, K., Tsunoda, T., Nishioka, M., Fukuyama, K., Yamamoto, K., Mekada, E. and Sasazuki, T. (2000). Involvement of deregulated epiregulin expression in tumorigenesis in vivo through activated Ki-Ras signaling pathway in human colon cancer cells. *Cancer Res* 60(24):6886–6889.

[62] Tørring, N., Jørgensen, P.E., Sørensen, B.S. and Nexø, E. (2000). Increased expression of heparin binding EGF (HB-EGF), amphiregulin, TGF alpha and epiregulin in androgen-independent prostate cancer cell lines. *Anticancer Res* 20(1A):91–95.

[63] Zhu, Z., Kleeff, J., Friess, H., Wang, L., Zimmermann, A., Yarden, Y., Büchler, M.W. and Korc, M. (2000). Epiregulin is Up-regulated in pancreatic cancer and stimulates pancreatic cancer cell growth. *Biochem Biophys Res Commun* 273(3):1019–1024.

[64] Waddington (1939). An Introduction to Modern Genetics. London: George Allen and Unwin.

[65] Bormann, F. (2008), Development of a Mulitplex Assay for robust Determination of epigenetic Changes in genomic Regions, associated with the Formation of Colorectal Cancer. Diploma thesis, Freie Universität Berlin.

[66] Watson, J.D. and Crick, F.H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 171(4356):737–738.

[67] Russo, V., Martienssen, R. and Riggs, A. (1996). Epigenetic Mechanisms of Gene Regulation. Cold Spring Harbor Laboratory Press.

[68] Bird, A. (2007). Perceptions of epigenetics. *Nature* 447(7143):396–398.

[69] Jones, P.A. and Baylin, S.B. (2002). The fundamental role of epigenetic events in cancer. *Nat Rev Genet* 3(6):415–428.

[70] Hotchkiss, R.D. (1948). The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. *J Biol Chem* 175(1):315–332.

[71] Roy, P.H. and Weissbach, A. (1975). DNA methylase from HeLa cell nuclei. *Nucleic Acids Res* 2(10):1669–1684.

[72] Yen, R.W., Vertino, P.M., Nelkin, B.D., Yu, J.J., el Deiry, W., Cumaraswamy, A., Lennon, G.G., Trask, B.J., Celano, P. and Baylin, S.B. (1992). Isolation and characterization of the cDNA encoding human DNA methyltransferase. *Nucleic Acids Res* 20(9):2287–2291.

[73] Lei, H., Oh, S.P., Okano, M., Jüttermann, R., Goss, K.A., Jaenisch, R. and Li, E. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* 122(10):3195–3205.

[74] Okano, M., Xie, S. and Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat Genet* 19(3):219–220.

[75] Hsieh, C.L. (1994). Dependence of transcriptional repression on CpG methylation density. *Mol Cell Biol* 14(8):5487–5494.

[76] Gardiner-Garden, M. and Frommer, M. (1987). CpG islands in vertebrate genomes. *J Mol Biol* 196(2):261–282.

[77] Li, E., Beard, C. and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. *Nature* 366(6453):362–365.

[78] Heard, E., Clerc, P. and Avner, P. (1997). X-chromosome inactivation in mammals. *Annu Rev Genet* 31:571–610.

[79] Brenet, F., Moh, M., Funk, P., Feierstein, E., Viale, A.J., Socci, N.D. and Scandura, J.M. (2011). DNA methylation of the first exon is tightly linked to transcriptional silencing. *PLoS One* 6(1):e14524.

[80] Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B. et al. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454(7205):766–770.

[81] Branco, M.R., Ficz, G. and Reik, W. (2012). Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat Rev Genet* 13(1):7–13.

[82] Kossel, A. (1928). The protamines and histones. London, New York [etc.] Longmans, Green and co.

[83] Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* 389(6648):251–260.

[84] Grant, P.A. (2001). A tale of histone modifications. *Genome Biol* 2(4):REVIEWS0003.

[85] Turner, B.M., Birley, A.J. and Lavender, J. (1992). Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophila polytene nuclei. *Cell* 69(2):375–384.

[86] Braunstein, M., Sobel, R.E., Allis, C.D., Turner, B.M. and Broach, J.R. (1996). Efficient transcriptional silencing in Saccharomyces cerevisiae requires a heterochromatin histone acetylation pattern. *Mol Cell Biol* 16(8):4349–4356.

[87] Allfrey, V.G., Faulkner, R. and Mirsky, A.E. (1964). Acetylation and Methylation of Histones and their possible role in the regulation of RNA-synthesis. *Proc Natl Acad Sci U S A* 51:786–794.

[88] Kuo, M.H., Brownell, J.E., Sobel, R.E., Ranalli, T.A., Cook, R.G., Edmondson, D.G., Roth, S.Y. and Allis, C.D. (1996). Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines. *Nature* 383(6597):269–272.

[89] Grant, P.A. and Berger, S.L. (1999). Histone acetyltransferase complexes. *Semin Cell Dev Biol* 10(2):169–177.

[90] Ekwall, K., Olsson, T., Turner, B.M., Cranston, G. and Allshire, R.C. (1997). Transient inhibition of histone deacetylation alters the structural and functional imprint at fission yeast centromeres. *Cell* 91(7):1021–1032.

[91] Nowak, S.J. and Corces, V.G. (2000). Phosphorylation of histone H3 correlates with transcriptionally active loci. *Genes Dev* 14(23):3003–3013.

[92] Cheung, P., Tanner, K.G., Cheung, W.L., Sassone-Corsi, P., Denu, J.M. and Allis, C.D. (2000). Synergistic coupling of histone H3 phosphorylation and acetylation in response to epidermal growth factor stimulation. *Mol Cell* 5(6):905–915.

[93] Rosenfeld, J.A., Wang, Z., Schones, D.E., Zhao, K., DeSalle, R. and Zhang, M.Q. (2009). Determination of enriched histone modifications in non-genic portions of the human genome. *BMC Genomics* 10:143.

[94] Gupta, S., Kim, S.Y., Artis, S., Molfese, D.L., Schumacher, A., Sweatt, J.D., Paylor, R.E. and Lubin, F.D. (2010). Histone methylation regulates memory formation. *J Neurosci* 30(10):3589–3599.

[95] Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410(6824):116–120.

[96] Comb, M. and Goodman, H.M. (1990). CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2. *Nucleic Acids Res* 18(13):3975–3982.

[97] Clark, S.J., Harrison, J. and Molloy, P.L. (1997). Sp1 binding is inhibited by (m)Cp(m)CpG methylation. *Gene* 195(1):67–71.

[98] Meehan, R.R., Lewis, J.D., McKay, S., Kleiner, E.L. and Bird, A.P. (1989). Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell* 58(3):499–507.

[99] Bird, A.P. and Wolffe, A.P. (1999). Methylation-induced repression–belts, braces, and chromatin. *Cell* 99(5):451–454.

[100] Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N. and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 393(6683):386–389.

[101] Ng, H.H., Zhang, Y., Hendrich, B., Johnson, C.A., Turner, B.M., Erdjument-Bromage, H., Tempst, P., Reinberg, D. and Bird, A. (1999). MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat Genet* 23(1):58–61.

[102] Jones, P.L., Veenstra, G.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J. and Wolffe, A.P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat Genet* 19(2):187–191.

[103] Ng, H.H., Jeppesen, P. and Bird, A. (2000). Active repression of methylated genes by the chromosomal protein MBD1. *Mol Cell Biol* 20(4):1394–1406.

[104] Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S. and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* 298(5595):1039–1043.

[105] Viré, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Eynde, A.V., Bernard, D., Vanderwinden, J.M. et al. (2006). The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* 439(7078):871–874.

[106] Daniel, J.M. and Reynolds, A.B. (1999). The catenin p120(ctn) interacts with Kaiso, a novel BTB/POZ domain zinc finger transcription factor. *Mol Cell Biol* 19(5):3614–3623.

[107] Filion, G.J.P., Zhenilo, S., Salozhin, S., Yamada, D., Prokhortchouk, E. and Defossez, P.A. (2006). A family of human zinc finger proteins that bind methylated DNA and repress transcription. *Mol Cell Biol* 26(1):169–181.

[108] Yoon, H.G., Chan, D.W., Reynolds, A.B., Qin, J. and Wong, J. (2003). N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein Kaiso. *Mol Cell* 12(3):723–734.

[109] Blattler, A., Yao, L., Wang, Y., Ye, Z., Jin, V.X. and Farnham, P.J. (2013). ZBTB33 binds unmethylated regions of the genome associated with actively expressed genes. *Epigenetics Chromatin* 6(1):13.

[110] Consortium, E.N.C.O.D.E.P., Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C.A., Doyle, F., Epstein, C.B., Frietze, S., Harrow, J. et al. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489(7414):57–74.

[111] Renda, M., Baglivo, I., Burgess-Beusse, B., Esposito, S., Fattorusso, R., Felsenfeld, G. and Pedone, P.V. (2007). Critical DNA binding interactions of the insulator protein CTCF: a small number of zinc fingers mediate strong binding, and a single finger-DNA interaction controls binding at imprinted loci. *J Biol Chem* 282(46):33336–33345.

[112] Filippova, G.N., Thienes, C.P., Penn, B.H., Cho, D.H., Hu, Y.J., Moore, J.M., Klesert, T.R., Lobanenkov, V.V. and Tapscott, S.J. (2001). CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat Genet* 28(4):335–343.

[113] Filippova, G.N., Fagerlie, S., Klenova, E.M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P.E., Collins, S.J. and Lobanenkov, V.V. (1996). An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Mol Cell Biol* 16(6):2802–2813.

[114] Vostrov, A.A. and Quitschke, W.W. (1997). The zinc finger protein CTCF binds to the APBbeta domain of the amyloid beta-protein precursor promoter. Evidence for a role in transcriptional activation. *J Biol Chem* 272(52):33353–33359.

[115] Bell, A.C., West, A.G. and Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98(3):387–396.

[116] Loukinov, D.I., Pugacheva, E., Vatolin, S., Pack, S.D., Moon, H., Chernukhin, I., Mannan, P., Larsson, E., Kanduri, C., Vostrov, A.A. et al. (2002). BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma. *Proc Natl Acad Sci U S A* 99(10):6806–6811.

[117] Jones, T.A., Ogunkolade, B.W., Szary, J., Aarum, J., Mumin, M.A., Patel, S., Pieri, C.A. and Sheer, D. (2011). Widespread expression of BORIS/CTCFL in normal and cancer cells. *PLoS One* 6(7):e22399.

[118] Klenova, E.M., Morse, H.C., Ohlsson, R. and Lobanenkov, V.V. (2002). The novel BORIS + CTCF gene family is uniquely involved in the epigenetics of normal biology and cancer. *Semin Cancer Biol* 12(5):399–414.

[119] Feinberg, A.P. and Vogelstein, B. (1983). Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* 301(5895):89–92.

[120] Gama-Sosa, M.A., Slagel, V.A., Trewyn, R.W., Oxenhandler, R., Kuo, K.C., Gehrke, C.W. and Ehrlich, M. (1983). The 5-methylcytosine content of DNA from human tumors. *Nucleic Acids Res* 11(19):6883–6894.

[121] Adorján, P., Distler, J., Lipscher, E., Model, F., Müller, J., Pelet, C., Braun, A., Florl, A.R., Gütig, D., Grabs, G. et al. (2002). Tumour class prediction and discovery by microarray-based DNA methylation analysis. *Nucleic Acids Res* 30(5):e21.

[122] Weber, M., Davies, J.J., Wittig, D., Oakeley, E.J., Haase, M., Lam, W.L. and Schübeler, D. (2005). Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat Genet* 37(8):853–862.

[123] Feinberg, A.P. and Vogelstein, B. (1983). Hypomethylation of ras oncogenes in primary human cancers. *Biochem Biophys Res Commun* 111(1):47–54.

[124] Wu, H., Chen, Y., Liang, J., Shi, B., Wu, G., Zhang, Y., Wang, D., Li, R., Yi, X., Zhang, H. et al. (2005). Hypomethylation-linked activation of PAX2 mediates tamoxifen-stimulated endometrial carcinogenesis. *Nature* 438(7070):981–987.

[125] Eden, A., Gaudet, F., Waghmare, A. and Jaenisch, R. (2003). Chromosomal instability and tumors promoted by DNA hypomethylation. *Science* 300(5618):455.

[126] Karpf, A.R. and ichi Matsui, S. (2005). Genetic disruption of cytosine DNA methyltransferase enzymes induces chromosomal instability in human cancer cells. *Cancer Res* 65(19):8635–8639.

[127] Suter, C.M., Martin, D.I. and Ward, R.L. (2004). Hypomethylation of L1 retrotransposons in colorectal cancer and adjacent normal tissue. *Int J Colorectal Dis* 19(2):95–101.

[128] Bestor, T.H. (2005). Transposons reanimated in mice. *Cell* 122(3):322–325.

[129] Cui, H., Cruz-Correa, M., Giardiello, F.M., Hutcheon, D.F., Kafonek, D.R., Brandenburg, S., Wu, Y., He, X., Powe, N.R. and Feinberg, A.P. (2003). Loss of IGF2 imprinting: a potential marker of colorectal cancer risk. *Science* 299(5613):1753–1755.

[130] Kaneda, A. and Feinberg, A.P. (2005). Loss of imprinting of IGF2: a common epigenetic modifier of intestinal tumor risk. *Cancer Res* 65(24):11236–11240.

[131] Fraga, M.F., Herranz, M., Espada, J., Ballestar, E., Paz, M.F., Ropero, S., Erkek, E., Bozdogan, O., Peinado, H., Niveleau, A. et al. (2004). A mouse skin multistage carcinogenesis model reflects the aberrant DNA methylation patterns of human tumors. *Cancer Res* 64(16):5527–5534.

[132] Esteller, M., Silva, J.M., Dominguez, G., Bonilla, F., Matias-Guiu, X., Lerma, E., Bussaglia, E., Prat, J., Harkes, I.C., Repasky, E.A. et al. (2000). Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. *J Natl Cancer Inst* 92(7):564–569.

[133] Gonzalez-Zulueta, M., Bender, C.M., Yang, A.S., Nguyen, T., Beart, R.W., Tornout, J.M.V. and Jones, P.A. (1995). Methylation of the 5' CpG island of the p16/CDKN2 tumor suppressor gene in normal and transformed human tissues correlates with gene silencing. *Cancer Res* 55(20):4531–4535.

[134] Herman, J.G., Merlo, A., Mao, L., Lapidus, R.G., Issa, J.P., Davidson, N.E., Sidransky, D. and Baylin, S.B. (1995). Inactivation of the CDKN2/p16/MTS1 gene is frequently associated with aberrant DNA methylation in all common human cancers. *Cancer Res* 55(20):4525–4530.

[135] Graff, J.R., Herman, J.G., Lapidus, R.G., Chopra, H., Xu, R., Jarrard, D.F., Isaacs, W.B., Pitha, P.M., Davidson, N.E. and Baylin, S.B. (1995). E-cadherin expression is silenced by DNA hypermethylation in human breast and prostate carcinomas. *Cancer Res* 55(22):5195–5199.

[136] Herman, J.G., Latif, F., Weng, Y., Lerman, M.I., Zbar, B., Liu, S., Samid, D., Duan, D.S., Gnarra, J.R. and Linehan, W.M. (1994). Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. *Proc Natl Acad Sci U S A* 91(21):9700–9704.

[137] Costello, J.F., Frühwald, M.C., Smiraglia, D.J., Rush, L.J., Robertson, G.P., Gao, X., Wright, F.A., Feramisco, J.D., Peltomäki, P., Lang, J.C. et al. (2000). Aberrant CpG-island methylation has non-random and tumour-type-specific patterns. *Nat Genet* 24(2):132–138.

[138] Esteller, M., Corn, P.G., Baylin, S.B. and Herman, J.G. (2001). A gene hypermethylation profile of human cancer. *Cancer Res* 61(8):3225–3229.

[139] Toyota, M., Ahuja, N., Ohe-Toyota, M., Herman, J.G., Baylin, S.B. and Issa, J.P. (1999). CpG island methylator phenotype in colorectal cancer. *Proc Natl Acad Sci U S A* 96(15):8681–8686.

[140] Issa, J.P.J., Shen, L. and Toyota, M. (2005). CIMP, at last. *Gastroenterology* 129(3):1121–1124.

[141] Samowitz, W.S., Albertsen, H., Herrick, J., Levin, T.R., Sweeney, C., Murtaugh, M.A., Wolff, R.K. and Slattery, M.L. (2005). Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer. *Gastroenterology* 129(3):837–845.

[142] Yamashita, K., Dai, T., Dai, Y., Yamamoto, F. and Perucho, M. (2003). Genetics supersedes epigenetics in colon cancer phenotype. *Cancer Cell* 4(2):121–131.

[143] Anacleto, C., Leopoldino, A.M., Rossi, B., Soares, F.A., Lopes, A., Rocha, J.C.C., Caballero, O., Camargo, A.A., Simpson, A.J.G. and Pena, S.D.J. (2005). Colorectal cancer "methylator phenotype": fact or artifact? *Neoplasia* 7(4):331–335.

[144] Sakai, T., Toguchida, J., Ohtani, N., Yandell, D.W., Rapaport, J.M. and Dryja, T.P. (1991). Allele-specific hypermethylation of the retinoblastoma tumor-suppressor gene. *Am J Hum Genet* 48(5):880–888.

[145] Ohtani-Fujita, N., Fujita, T., Aoike, A., Osifchin, N.E., Robbins, P.D. and Sakai, T. (1993). CpG methylation inactivates the promoter activity of the human retinoblastoma tumor-suppressor gene. *Oncogene* 8(4):1063–1067.

[146] Dunaief, J.L., Strober, B.E., Guha, S., Khavari, P.A., Alin, K., Luban, J., Begemann, M., Crabtree, G.R. and Goff, S.P. (1994). The retinoblastoma protein and BRG1 form a complex and cooperate to induce cell cycle arrest. *Cell* 79(1):119–130.

[147] Luo, R.X., Postigo, A.A. and Dean, D.C. (1998). Rb interacts with histone deacetylase to repress transcription. *Cell* 92(4):463–473.

[148] Dahiya, A., Wong, S., Gonzalo, S., Gavin, M. and Dean, D.C. (2001). Linking the Rb and polycomb pathways. *Mol Cell* 8(3):557–569.

[149] Robertson, K.D., Ait-Si-Ali, S., Yokochi, T., Wade, P.A., Jones, P.L. and Wolffe, A.P. (2000). DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nat Genet* 25(3):338–342.

[150] Obeyesekere, M.N., Herbert, J.R. and Zimmerman, S.O. (1995). A model of the G1 phase of the cell cycle incorporating cyclin E/cdk2 complex and retinoblastoma protein. *Oncogene* 11(6):1199–1205.

[151] Fraga, M.F., Ballestar, E., Villar-Garea, A., Boix-Chornet, M., Espada, J., Schotta, G., Bonaldi, T., Haydon, C., Ropero, S., Petrie, K. et al. (2005). Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat Genet* 37(4):391–400.

[152] Ozdağ, H., Teschendorff, A.E., Ahmed, A.A., Hyland, S.J., Blenkiron, C., Bobrow, L., Veerakumarasivam, A., Burtt, G., Subkhankulova, T., Arends, M.J. et al. (2006). Differential expression of selected histone modifier genes in human solid cancers. *BMC Genomics* 7:90.

[153] Siewert, J. and Lordick, F. (2006). Standards in der interdisziplinären Diagnostik und Therapie. In Kompendium Internistische Onkologie, (eds.) H.J. Schmoll, K. Höffken and K. Possinger, pp. 523–540, Springer Berlin Heidelberg.

[154] Molls, M. and Röper, B. (2006). Strahlenbiologische Grundlagen der Strahlentherapie. In Kompendium Internistische Onkologie, (eds.) H.J. Schmoll, K. Höffken and K. Possinger, pp. 541–549, Springer Berlin Heidelberg.

[155] Pfreundschuh, M. (2006). Prinzipien der medikamentösen Tumortherapie. In Kompendium Internistische Onkologie, (eds.) H.J. Schmoll, K. Höffken and K. Possinger, pp. 651–702, Springer Berlin Heidelberg.

[156] André, T., Louvet, C., Maindrault-Goebel, F., Couteau, C., Mabro, M., Lotz, J.P., Gilles-Amar, V., Krulik, M., Carola, E., Izrael, V. et al. (1999). CPT-11 (irinotecan) addition to bimonthly, high-dose leucovorin and bolus and continuous-infusion 5-fluorouracil (FOLFIRI) for pretreated metastatic colorectal cancer. GERCOR. *Eur J Cancer* 35(9):1343–1347.

[157] Gill, G.N., Kawamoto, T., Cochet, C., Le, A., Sato, J.D., Masui, H., McLeod, C. and Mendelsohn, J. (1984). Monoclonal anti-epidermal growth factor receptor antibodies which are inhibitors of epidermal growth factor binding and antagonists of epidermal growth factor binding and antagonists of epidermal growth factor-stimulated tyrosine protein kinase activity. *J Biol Chem* 259(12):7755–7760.

[158] Arteaga, C.L., Coronado, E. and Osborne, C.K. (1988). Blockade of the epidermal growth factor receptor inhibits transforming growth factor alpha-induced but not estrogen-induced growth of hormone-dependent human breast cancer. *Mol Endocrinol* 2(11):1064–1069.

[159] Ennis, B.W., Valverius, E.M., Bates, S.E., Lippman, M.E., Bellot, F., Kris, R., Schlessinger, J., Masui, H., Goldenberg, A. and Mendelsohn, J. (1989). Anti-epidermal growth factor receptor antibodies inhibit the autocrine-stimulated growth of MDA-468 human breast cancer cells. *Mol Endocrinol* 3(11):1830–1838.

[160] Karnes, W.E., Walsh, J.H., Wu, S.V., Kim, R.S., Martin, M.G., Wong, H.C., Mendelsohn, J., Park, J.G. and Cuttitta, F. (1992). Autonomous proliferation of colon cancer cells that coexpress transforming growth factor alpha and its receptor. Variable effects of receptor-blocking antibody. *Gastroenterology* 102(2):474–485.

[161] Wu, X., Fan, Z., Masui, H., Rosen, N. and Mendelsohn, J. (1995). Apoptosis induced by an anti-epidermal growth factor receptor monoclonal antibody in a human colorectal carcinoma cell line and its delay by insulin. *J Clin Invest* 95(4):1897–1905.

[162] Atlas, I., Mendelsohn, J., Baselga, J., Fair, W.R., Masui, H. and Kumar, R. (1992). Growth regulation of human renal carcinoma cells: role of transforming growth factor alpha. *Cancer Res* 52(12):3335–3339.

[163] Peng, D., Fan, Z., Lu, Y., DeBlasio, T., Scher, H. and Mendelsohn, J. (1996). Anti-epidermal growth factor receptor monoclonal antibody 225 up-regulates p27KIP1 and induces G1 arrest in prostatic cancer cell line DU145. *Cancer Res* 56(16):3666–3669.

[164] Hofer, D.R., Sherwood, E.R., Bromberg, W.D., Mendelsohn, J., Lee, C. and Kozlowski, J.M. (1991). Autonomous growth of androgen-independent human prostatic carcinoma cells: role of transforming growth factor alpha. *Cancer Res* 51(11):2780–2785.

[165] Masui, H., Kawamoto, T., Sato, J.D., Wolf, B., Sato, G. and Mendelsohn, J. (1984). Growth inhibition of human tumor cells in athymic mice by anti-epidermal growth factor receptor monoclonal antibodies. *Cancer Res* 44(3):1002–1007.

[166] Masui, H., Moroyama, T. and Mendelsohn, J. (1986). Mechanism of antitumor activity in mice for anti-epidermal growth factor receptor monoclonal antibodies with different isotypes. *Cancer Res* 46(11):5592–5598.

[167] Mendelsohn, J. (1989). Anti-EGF receptor monoclonal antibodies: biological studies and potential clinical applications. *Trans Am Clin Climatol Assoc* 100:31–38.

[168] Divgi, C.R., Welt, S., Kris, M., Real, F.X., Yeh, S.D., Gralla, R., Merchant, B., Schweighart, S., Unger, M. and Larson, S.M. (1991). Phase I and imaging trial of indium 111-labeled anti-epidermal growth factor receptor monoclonal antibody 225 in patients with squamous cell lung carcinoma. *J Natl Cancer Inst* 83(2):97–104.

[169] Naramura, M., Gillies, S.D., Mendelsohn, J., Reisfeld, R.A. and Mueller, B.M. (1993). Therapeutic potential of chimeric and murine anti-(epidermal growth factor receptor) antibodies in a metastasis model for human melanoma. *Cancer Immunol Immunother* 37(5):343–349.

[170] Schroff, R.W., Foon, K.A., Beatty, S.M., Oldham, R.K. and Morgan, A.C. (1985). Human anti-murine immunoglobulin responses in patients receiving monoclonal antibody therapy. *Cancer Res* 45(2):879–885.

[171] Huang, S.M., Bock, J.M. and Harari, P.M. (1999). Epidermal growth factor receptor blockade with C225 modulates proliferation, apoptosis, and radiosensitivity in squamous cell carcinomas of the head and neck. *Cancer Res* 59(8):1935–1940.

[172] Ciardiello, F., Caputo, R., Bianco, R., Damiano, V., Pomatico, G., Pepe, S., Bianco, A.R., Agrawal, S., Mendelsohn, J. and Tortora, G. (1998). Cooperative inhibition of renal cancer growth by anti-epidermal growth factor receptor antibody and protein kinase A antisense oligonucleotide. *J Natl Cancer Inst* 90(14):1087–1094.

[173] Prewett, M., Rothman, M., Waksal, H., Feldman, M., Bander, N.H. and Hicklin, D.J. (1998). Mouse-human chimeric anti-epidermal growth factor receptor antibody C225 inhibits the growth of human renal cell carcinoma xenografts in nude mice. *Clin Cancer Res* 4(12):2957–2966.

[174] Ciardiello, F., Bianco, R., Damiano, V., Lorenzo, S.D., Pepe, S., Placido, S.D., Fan, Z., Mendelsohn, J., Bianco, A.R. and Tortora, G. (1999). Antitumor activity of sequential treatment with topotecan and anti-epidermal growth factor receptor monoclonal antibody C225. *Clin Cancer Res* 5(4):909–916.

[175] Cutsem, E.V., Köhne, C.H., Hitre, E., Zaluski, J., Chien, C.R.C., Makhson, A., D'Haens, G., Pintér, T., Lim, R., Bodoky, G. et al. (2009). Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. *N Engl J Med* 360(14):1408–1417.

[176] Moyer, J.D., Barbacci, E.G., Iwata, K.K., Arnold, L., Boman, B., Cunningham, A., DiOrio, C., Doty, J., Morin, M.J., Moyer, M.P. et al. (1997). Induction of apoptosis and cell cycle arrest by CP-358,774, an inhibitor of epidermal growth factor receptor tyrosine kinase. *Cancer Res* 57(21):4838–4848.

[177] Albanell, J., Codony-Servat, J., Rojo, F., Campo, J.M.D., Sauleda, S., Anido, J., Raspall, G., Giralt, J., Roselló, J., Nicholson, R.I. et al. (2001). Activated extracellular signal-regulated kinases: association with epidermal growth factor receptor/transforming growth factor alpha expression in head and neck squamous carcinoma and inhibition by anti-epidermal growth factor receptor treatments. *Cancer Res* 61(17):6500–6510.

[178] Giaccone, G., Herbst, R.S., Manegold, C., Scagliotti, G., Rosell, R., Miller, V., Natale, R.B., Schiller, J.H., Pawel, J.V., Pluzanska, A. et al. (2004). Gefitinib in combination with gemcitabine and cisplatin in advanced non-small-cell lung cancer: a phase III trial–INTACT 1. *J Clin Oncol* 22(5):777–784.

[179] Herbst, R.S., Giaccone, G., Schiller, J.H., Natale, R.B., Miller, V., Manegold, C., Scagliotti, G., Rosell, R., Oliff, I., Reeves, J.A. et al. (2004). Gefitinib in combination with paclitaxel and carboplatin in advanced non-small-cell lung cancer: a phase III trial–INTACT 2. *J Clin Oncol* 22(5):785–794.

[180] Herbst, R.S., Prager, D., Hermann, R., Fehrenbacher, L., Johnson, B.E., Sandler, A., Kris, M.G., Tran, H.T., Klein, P., Li, X. et al. (2005). TRIBUTE: a phase III trial of erlotinib hydrochloride (OSI-774) combined with carboplatin and paclitaxel chemotherapy in advanced non-small-cell lung cancer. *J Clin Oncol* 23(25):5892–5899.

[181] Gatzemeier, U., Pluzanska, A., Szczesna, A., Kaukel, E., Roubec, J., Rosa, F.D., Milanowski, J., Karnicka-Mlodkowski, H., Pesek, M., Serwatowski, P. et al. (2007). Phase III study of erlotinib in combination with cisplatin and gemcitabine in advanced non-small-cell lung cancer: the Tarceva Lung Cancer Investigation Trial. *J Clin Oncol* 25(12):1545–1552.

[182] Shepherd, F.A., Pereira, J.R., Ciuleanu, T., Tan, E.H., Hirsh, V., Thongprasert, S., Campos, D., Maoleekoonpiroj, S., Smylie, M., Martins, R. et al. (2005). Erlotinib in previously treated non-small-cell lung cancer. *N Engl J Med* 353(2):123–132.

[183] Mok, T.S., Wu, Y.L., Thongprasert, S., Yang, C.H., Chu, D.T., Saijo, N., Sunpaweravong, P., Han, B., Margono, B., Ichinose, Y. et al. (2009). Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 361(10):947–957.

[184] Neal, J.W. (2010). The SATURN trial: the value of maintenance erlotinib in patients with non-small-cell lung cancer. *Future Oncol* 6(12):1827–1832.

[185] Townsley, C.A., Major, P., Siu, L.L., Dancey, J., Chen, E., Pond, G.R., Nicklee, T., Ho, J., Hedley, D., Tsao, M. et al. (2006). Phase II study of erlotinib (OSI-774) in patients with metastatic colorectal cancer. *Br J Cancer* 94(8):1136–1143.

[186] Meyerhardt, J.A., Stuart, K., Fuchs, C.S., Zhu, A.X., Earle, C.C., Bhargava, P., Blaszkowsky, L., Enzinger, P., Mayer, R.J., Battu, S. et al. (2007). Phase II study of FOLFOX, bevacizumab and erlotinib as first-line therapy for patients with metastatic colorectal cancer. *Ann Oncol* 18(7):1185–1189.

[187] Santoro, A., Comandone, A., Rimassa, L., Granetti, C., Lorusso, V., Oliva, C., Ronzoni, M., Siena, S., Zuradelli, M., Mari, E. et al. (2008). A phase II randomized multicenter trial of gefitinib plus FOLFIRI and FOLFIRI alone in patients with metastatic colorectal cancer. *Ann Oncol* 19(11):1888–1893.

[188] Weickhardt, A.J., Price, T.J., Chong, G., Gebski, V., Pavlakis, N., Johns, T.G., Azad, A., Skrinos, E., Fluck, K., Dobrovic, A. et al. (2012). Dual targeting of the epidermal growth factor receptor using the combination of cetuximab and erlotinib: preclinical evaluation and results of the phase II DUX study in chemotherapy-refractory, advanced colorectal cancer. *J Clin Oncol* 30(13):1505–1512.

[189] Ishikawa, N., Daigo, Y., Takano, A., Taniwaki, M., Kato, T., Hayama, S., Murakami, H., Takeshima, Y., Inai, K., Nishimura, H. et al. (2005). Increases of amphiregulin and transforming growth factor-alpha in serum as predictors of poor response to gefitinib among patients with advanced non-small cell lung cancers. *Cancer Res* 65(20):9176–9184.

[190] Masago, K., Fujita, S., Hatachi, Y., Fukuhara, A., Sakuma, K., Ichikawa, M., Kim, Y.H., Mio, T. and Mishima, M. (2008). Clinical significance of pretreatment serum amphiregulin and transforming growth factor-alpha, and an epidermal growth factor receptor somatic mutation in patients with advanced non-squamous, non-small cell lung cancer. *Cancer Sci* 99(11):2295–2301.

[191] Yonesaka, K., Zejnullahu, K., Lindeman, N., Homes, A.J., Jackman, D.M., Zhao, F., Rogers, A.M., Johnson, B.E. and Jänne, P.A. (2008). Autocrine production of amphiregulin predicts sensitivity to both gefitinib and cetuximab in EGFR wild-type cancers. *Clin Cancer Res* 14(21):6963–6973.

[192] Vollebergh, M.A., Kappers, I., Klomp, H.M., Buning-Kager, J.C., Korse, C.M., Hauptmann, M., de Visser, K.E., van den Heuvel, M.M. and Linn, S.C. (2010). Ligands of epidermal growth factor receptor and the insulin-like growth factor family as serum biomarkers for response to epidermal growth factor receptor inhibitors in patients with advanced non-small cell lung cancer. *J Thorac Oncol* 5(12):1939–1948.

[193] Khambata-Ford, S., Garrett, C.R., Meropol, N.J., Basik, M., Harbison, C.T., Wu, S., Wong, T.W., Huang, X., Takimoto, C.H., Godwin, A.K. et al. (2007). Expression of epiregulin and amphiregulin and K-ras mutation status predict disease control in metastatic colorectal cancer patients treated with cetuximab. *J Clin Oncol* 25(22):3230–3237.

[194] Jacobs, B., Roock, W.D., Piessevaux, H., Oirbeek, R.V., Biesmans, B., Schutter, J.D., Fieuws, S., Vandesompele, J., Peeters, M., Laethem, J.L.V. et al. (2009). Amphiregulin and epiregulin mRNA expression in primary tumors predicts outcome in metastatic colorectal cancer treated with cetuximab. *J Clin Oncol* 27(30):5068–5074.

[195] Pentheroudakis, G., Kotoula, V., Roock, W.D., Kouvatseas, G., Papakostas, P., Makatsoris, T., Papamichael, D., Xanthakis, I., Sgouros, J., Televantou, D. et al. (2013). Biomarkers of

benefit from cetuximab-based therapy in metastatic colorectal cancer: interaction of EGFR ligand expression with RAS/RAF, PIK3CA genotypes. *BMC Cancer* 13(1):49.

[196] Yoshida, M., Shimura, T., Sato, M., Ebi, M., Nakazawa, T., Takeyama, H. and Joh, T. (2013). A novel predictive strategy by immunohistochemical analysis of four EGFR ligands in metastatic colorectal cancer treated with anti-EGFR antibodies. *J Cancer Res Clin Oncol* 139(3):367–378.

[197] Lee, W.H., Morton, R.A., Epstein, J.I., Brooks, J.D., Campbell, P.A., Bova, G.S., Hsieh, W.S., Isaacs, W.B. and Nelson, W.G. (1994). Cytidine methylation of regulatory sequences near the pi-class glutathione S-transferase gene accompanies human prostatic carcinogenesis. *Proc Natl Acad Sci U S A* 91(24):11733–11737.

[198] Jerónimo, C., Usadel, H., Henrique, R., Oliveira, J., Lopes, C., Nelson, W.G. and Sidransky, D. (2001). Quantitation of GSTP1 methylation in non-neoplastic prostatic tissue and organ-confined prostate adenocarcinoma. *J Natl Cancer Inst* 93(22):1747–1752.

[199] Esteller, M., Fraga, M.F., Guo, M., Garcia-Foncillas, J., Hedenfalk, I., Godwin, A.K., Trojan, J., Vaurs-Barrire, C., Bignon, Y.J., Ramus, S. et al. (2001). DNA methylation patterns in hereditary human cancers mimic sporadic tumorigenesis. *Hum Mol Genet* 10(26):3001–3007.

[200] Esteller, M., Sanchez-Cespedes, M., Rosell, R., Sidransky, D., Baylin, S.B. and Herman, J.G. (1999). Detection of aberrant promoter hypermethylation of tumor suppressor genes in serum DNA from non-small cell lung cancer patients. *Cancer Res* 59(1):67–70.

[201] Cairns, P., Esteller, M., Herman, J.G., Schoenberg, M., Jeronimo, C., Sanchez-Cespedes, M., Chow, N.H., Grasso, M., Wu, L., Westra, W.B. et al. (2001). Molecular detection of prostate cancer in urine by GSTP1 hypermethylation. *Clin Cancer Res* 7(9):2727–2730.

[202] Fackler, M.J., Malone, K., Zhang, Z., Schilling, E., Garrett-Mayer, E., Swift-Scanlan, T., Lange, J., Nayar, R., Davidson, N.E., Khan, S.A. et al. (2006). Quantitative multiplex methylation-specific PCR analysis doubles detection of tumor cells in breast ductal fluid. *Clin Cancer Res* 12(11 Pt 1):3306–3310.

[203] Roh, J.L., Wang, V.X., Manola, J., Sidransky, D. and Koch, W.M. (2013). Clinical Correlates of Promoter Hypermethylation of Four Target Genes in Head and Neck Cancer: A Cooperative Group Correlative Study. *Clin Cancer Res* .

[204] Mitomi, H., Fukui, N., Tanaka, N., Kanazawa, H., Saito, T., Matsuoka, T. and Yao, T. (2010). Aberrant p16((INK4a)) methylation is a frequent event in colorectal cancers: prognostic value and relation to mRNA expression and immunoreactivity. *J Cancer Res Clin Oncol* 136(2):323–331.

[205] Komine, C., Watanabe, T., Katayama, Y., Yoshino, A., Yokoyama, T. and Fukushima, T. (2003). Promoter hypermethylation of the DNA repair gene O6-methylguanine-DNA methyltransferase is an independent predictor of shortened progression free survival in patients with low-grade diffuse astrocytomas. *Brain Pathol* 13(2):176–184.

[206] Esteller, M., Gaidano, G., Goodman, S.N., Zagonel, V., Capello, D., Botto, B., Rossi, D., Gloghini, A., Vitolo, U., Carbone, A. et al. (2002). Hypermethylation of the DNA repair gene O(6)-methylguanine DNA methyltransferase and survival of patients with diffuse large B-cell lymphoma. *J Natl Cancer Inst* 94(1):26–32.

[207] Esteller, M., Garcia-Foncillas, J., Andion, E., Goodman, S.N., Hidalgo, O.F., Vanaclocha, V., Baylin, S.B. and Herman, J.G. (2000). Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *N Engl J Med* 343(19):1350–1354.

[208] Hegi, M.E., Diserens, A.C., Gorlia, T., Hamou, M.F., de Tribolet, N., Weller, M., Kros, J.M., Hainfellner, J.A., Mason, W., Mariani, L. et al. (2005). MGMT gene silencing and benefit from temozolomide in glioblastoma. *N Engl J Med* 352(10):997–1003.

[209] Strathdee, G., MacKean, M.J., Illand, M. and Brown, R. (1999). A role for methylation of the hMLH1 promoter in loss of hMLH1 expression and drug resistance in ovarian cancer. *Oncogene* 18(14):2335–2341.

[210] Ferreri, A.J.M., Dell'Oro, S., Capello, D., Ponzoni, M., Iuzzolino, P., Rossi, D., Pasini, F., Ambrosetti, A., Orvieto, E., Ferrarese, F. et al. (2004). Aberrant methylation in the promoter region of the reduced folate carrier gene is a potential mechanism of resistance to methotrexate in primary central nervous system lymphomas. *Br J Haematol* 126(5):657–664.

[211] Müller, C.I., Rüter, B., Koeffler, H.P. and Lübbert, M. (2006). DNA hypermethylation of myeloid cells, a novel therapeutic target in MDS and AML. *Curr Pharm Biotechnol* 7(5):315–321.

[212] Cheng, J.C., Matsen, C.B., Gonzales, F.A., Ye, W., Greer, S., Marquez, V.E., Jones, P.A. and Selker, E.U. (2003). Inhibition of DNA methylation and reactivation of silenced genes by zebularine. *J Natl Cancer Inst* 95(5):399–409.

[213] Marks, P.A. and Breslow, R. (2007). Dimethyl sulfoxide to vorinostat: development of this histone deacetylase inhibitor as an anticancer drug. *Nat Biotechnol* 25(1):84–90.

[214] Kim, H.J. and Bae, S.C. (2011). Histone deacetylase inhibitors: molecular mechanisms of action and clinical trials as anti-cancer drugs. *Am J Transl Res* 3(2):166–179.

[215] Garcia-Manero, G., Yang, H., Bueso-Ramos, C., Ferrajoli, A., Cortes, J., Wierda, W.G., Faderl, S., Koller, C., Morris, G., Rosner, G. et al. (2008). Phase 1 study of the histone deacetylase inhibitor vorinostat (suberoylanilide hydroxamic acid [SAHA]) in patients with advanced leukemias and myelodysplastic syndromes. *Blood* 111(3):1060–1066.

[216] Razak, A.R.A., Hotte, S.J., Siu, L.L., Chen, E.X., Hirte, H.W., Powers, J., Walsh, W., Stayner, L.A., Laughlin, A., Novotny-Diermayr, V. et al. (2011). Phase I clinical, pharmacokinetic and pharmacodynamic study of SB939, an oral histone deacetylase (HDAC) inhibitor, in patients with advanced solid tumours. *Br J Cancer* 104(5):756–762.

[217] Dong, M., Ning, Z.Q., Xing, P.Y., Xu, J.L., Cao, H.X., Dou, G.F., Meng, Z.Y., Shi, Y.K., Lu, X.P. and Feng, F.Y. (2012). Phase I study of chidamide (CS055/HBI-8000), a new histone deacetylase inhibitor, in patients with advanced solid tumors and lymphomas. *Cancer Chemother Pharmacol* 69(6):1413–1422.

[218] Chu, B.F., Karpenko, M.J., Liu, Z., Aimiuwu, J., Villalona-Calero, M.A., Chan, K.K., Grever, M.R. and Otterson, G.A. (2013). Phase I study of 5-aza-2'-deoxycytidine in combination with valproic acid in non-small-cell lung cancer. *Cancer Chemother Pharmacol* 71(1):115–121.

[219] Olsen, E.A., Kim, Y.H., Kuzel, T.M., Pacheco, T.R., Foss, F.M., Parker, S., Frankel, S.R., Chen, C., Ricker, J.L., Arduino, J.M. et al. (2007). Phase IIb multicenter trial of vorinostat in patients with persistent, progressive, or treatment refractory cutaneous T-cell lymphoma. *J Clin Oncol* 25(21):3109–3115.

[220] Blum, K.A., Advani, A., Fernandez, L., Jagt, R.V.D., Brandwein, J., Kambhampati, S., Kassis, J., Davis, M., Bonfils, C., Dubay, M. et al. (2009). Phase II study of the histone deacetylase inhibitor MGCD0103 in patients with previously treated chronic lymphocytic leukaemia. *Br J Haematol* 147(4):507–514.

[221] Galli, M., Salmoiraghi, S., Golay, J., Gozzini, A., Crippa, C., Pescosta, N. and Rambaldi, A. (2010). A phase II multiple dose clinical trial of histone deacetylase inhibitor ITF2357 in patients with relapsed or progressive multiple myeloma. *Ann Hematol* 89(2):185–190.

[222] Modesitt, S.C., Sill, M., Hoffman, J.S., Bender, D.P. and Group, G.O. (2008). A phase II study of vorinostat in the treatment of persistent or recurrent epithelial ovarian or primary peritoneal carcinoma: a Gynecologic Oncology Group study. *Gynecol Oncol* 109(2):182–186.

[223] Vansteenkiste, J., Cutsem, E.V., Dumez, H., Chen, C., Ricker, J.L., Randolph, S.S. and Schöffski, P. (2008). Early phase II trial of oral vorinostat in relapsed or refractory breast, colorectal, or non-small cell lung cancer. *Invest New Drugs* 26(5):483–488.

[224] Blumenschein, G.R., Kies, M.S., Papadimitrakopoulou, V.A., Lu, C., Kumar, A.J., Ricker, J.L., Chiao, J.H., Chen, C. and Frankel, S.R. (2008). Phase II trial of the histone deacetylase inhibitor vorinostat (Zolinza, suberoylanilide hydroxamic acid, SAHA) in patients with recurrent and/or metastatic head and neck cancer. *Invest New Drugs* 26(1):81–87.

[225] Nimmanapalli, R., Fuino, L., Stobaugh, C., Richon, V. and Bhalla, K. (2003). Cotreatment with the histone deacetylase inhibitor suberoylanilide hydroxamic acid (SAHA) enhances imatinib-induced apoptosis of Bcr-Abl-positive human acute leukemia cells. *Blood* 101(8):3236–3239.

[226] Rosato, R.R., Almenara, J.A., Yu, C. and Grant, S. (2004). Evidence of a functional role for p21WAF1/CIP1 down-regulation in synergistic antileukemic interactions between the histone deacetylase inhibitor sodium butyrate and flavopiridol. *Mol Pharmacol* 65(3):571–581.

[227] Almenara, J., Rosato, R. and Grant, S. (2002). Synergistic induction of mitochondrial damage and apoptosis in human leukemia cells by flavopiridol and the histone deacetylase inhibitor suberoylanilide hydroxamic acid (SAHA). *Leukemia* 16(7):1331–1343.

[228] Camphausen, K., Burgan, W., Cerra, M., Oswald, K.A., Trepel, J.B., Lee, M.J. and Tofilon, P.J. (2004). Enhanced radiation-induced cell killing and prolongation of gammaH2AX foci expression by the histone deacetylase inhibitor MS-275. *Cancer Res* 64(1):316–321.

[229] Rashid, S.F., Moore, J.S., Walker, E., Driver, P.M., Engel, J., Edwards, C.E., Brown, G., Uskokovic, M.R. and Campbell, M.J. (2001). Synergistic growth inhibition of prostate cancer cells by 1 alpha,25 Dihydroxyvitamin D(3) and its 19-nor-hexafluoride analogs in combination with either sodium butyrate or trichostatin A. *Oncogene* 20(15):1860–1872.

[230] Rozen, S. and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 132:365–386.

[231] Kasprzyk, A. (2011). BioMart: driving a paradigm change in biological data management. *Database (Oxford)* 2011:bar049.

[232] Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R. et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947–2948.

[233] Vincze, T., Posfai, J. and Roberts, R.J. (2003). NEBcutter: A program to cleave DNA with restriction enzymes. *Nucleic Acids Res* 31(13):3688–3691.

[234] R Core Team (2012). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0.

[235] Stothard, P. (2000). The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. *Biotechniques* 28(6):1102, 1104.

[236] Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M. and Haussler, D. (2002). The human genome browser at UCSC. *Genome Res* 12(6):996–1006.

[237] Rosenbloom, K.R., Sloan, C.A., Malladi, V.S., Dreszer, T.R., Learned, K., Kirkup, V.M., Wong, M.C., Maddren, M., Fang, R., Heitner, S.G. et al. (2013). ENCODE Data in the UCSC Genome Browser: year 5 update. *Nucleic Acids Res* 41(D1):D56–D63.

[238] Applied Biosystems (2008). Guide to Performing Relative Quantitation of Gene Expression Using Real-Time Quantitative PCR. 4371095 Rev B.

[239] Klug, M. and Rehli, M. (2006). Functional analysis of promoter CpG methylation using a CpG-free luciferase reporter vector. *Epigenetics* 1(3):127–130.

[240] Sers, C., Kuner, R., Falk, C.S., Lund, P., Sueltmann, H., Braun, M., Buness, A., Ruschhaupt, M., Conrad, J., Mang-Fatehi, S. et al. (2009). Down-regulation of HLA Class I and NKG2D ligands through a concerted action of MAPK and DNA methyltransferases in colorectal cancer cells. *Int J Cancer* 125(7):1626–1639.

[241] Bartels, S. (2012). Aktivierung der Signaltransduktion durch exosomale und freie EGFR Liganden in kolorektalen Karzinomzellen. Master's thesis, Charité, Berlin.

[242] Sasai, N., Nakao, M. and Defossez, P.A. (2010). Sequence-specific recognition of methylated DNA by human zinc-finger proteins. *Nucleic Acids Res* 38(15):5015–5022.

[243] Zhang, Q.C., Jiang, S.J., Zhang, S. and Ma, X.B. (2012). Histone deacetylase inhibitor trichostatin A enhances anti-tumor effects of docetaxel or erlotinib in A549 cell line. *Asian Pac J Cancer Prev* 13(7):3471–3476.

[244] Bianchi, L., Bruzzese, F., Leone, A., Gagliardi, A., Puglia, M., Gennaro, E.D., Rocco, M., Gimigliano, A., Pucci, B., Armini, A. et al. (2011). Proteomic analysis identifies differentially expressed proteins after HDAC vorinostat and EGFR inhibitor gefitinib treatments in Hep-2 cancer cells. *Proteomics* 11(18):3725–3742.

[245] Bruzzese, F., Leone, A., Rocco, M., Carbone, C., Piro, G., Caraglia, M., Gennaro, E.D. and Budillon, A. (2011). HDAC inhibitor vorinostat enhances the antitumor effect of gefitinib in squamous cell carcinoma of head and neck by modulating ErbB receptor expression and reverting EMT. *J Cell Physiol* 226(9):2378–2390.

[246] Jhawer, M., Goel, S., Wilson, A.J., Montagna, C., Ling, Y.H., Byun, D.S., Nasser, S., Arango, D., Shin, J., Klampfer, L. et al. (2008). PIK3CA mutation/PTEN expression status predicts response of colon cancer cells to the epidermal growth factor receptor inhibitor cetuximab. *Cancer Res* 68(6):1953–1961.

[247] Sørensen, B.S., Tørring, N., Bor, M.V. and Nexø, E. (2000). Quantitation of the mRNA expression of the epidermal growth factor system: selective induction of heparin-binding epidermal growth factor-like growth factor and amphiregulin expression by growth factor stimulation of prostate stromal cells. *J Lab Clin Med* 136(3):209–217.

[248] Tse, A.C.K. and Ge, W. (2009). Differential regulation of betacellulin and heparin-binding EGF-like growth factor in cultured zebrafish ovarian follicle cells by EGF family ligands. *Comp Biochem Physiol A Mol Integr Physiol* 153(1):13–17.

[249] Vallejo, G., Mestre-Citrinovitz, A.C., Mönckedieck, V., Grümmer, R., Winterhager, E. and Saragüeta, P. (2011). Ovarian steroid receptors and activated MAPK in the regional decidualization in rats. *Biol Reprod* 84(5):1063–1071.

[250] Sizemore, N., Cox, A.D., Barnard, J.A., Oldham, S.M., Reynolds, E.R., Der, C.J. and Coffey, R.J. (1999). Pharmacological inhibition of Ras-transformed epithelial cell growth is linked to down-regulation of epidermal growth factor-related peptides. *Gastroenterology* 117(3):567–576.

[251] Katoh, Y. and Katoh, M. (2006). Canonical WNT signaling pathway and human AREG. *Int J Mol Med* 17(6):1163–1166.

[252] Lüchtenborg, M., Weijenberg, M.P., Roemen, G.M.J.M., de Brune, A.P., van den Brandt, P.A., Lentjes, M.H.F.M., Brink, M., van Engeland, M., Goldbohm, R.A. and de Goeij, A.F.P.M. (2004). APC mutations in sporadic colorectal carcinomas from The Netherlands Cohort Study. *Carcinogenesis* 25(7):1219–1226.

[253] Morin, P.J., Sparks, A.B., Korinek, V., Barker, N., Clevers, H., Vogelstein, B. and Kinzler, K.W. (1997). Activation of beta-catenin-Tcf signaling in colon cancer by mutations in beta-catenin or APC. *Science* 275(5307):1787–1790.

[254] Mack, E.M., Smith, J.E., Kurz, S.G. and Wood, J.R. (2012). cAMP-dependent regulation of ovulatory response genes is amplified by IGF1 due to synergistic effects on Akt phosphorylation and NF-B transcription factors. *Reproduction* 144(5):595–602.

[255] Plowman, G.D., Green, J.M., McDonald, V.L., Neubauer, M.G., Disteche, C.M., Todaro, G.J. and Shoyab, M. (1990). The amphiregulin gene encodes a novel epidermal growth factor-related protein with tumor-inhibitory activity. *Mol Cell Biol* 10(5):1969–1981.

[256] Johansson, C.C., Yndestad, A., Enserink, J.M., Ree, A.H., Aukrust, P. and Taskén, K. (2004). The epidermal growth factor-like growth factor amphiregulin is strongly induced by the adenosine 3',5'-monophosphate pathway in various cell types. *Endocrinology* 145(11):5177–5184.

[257] Nakajima, T., Uchida, C., Anderson, S.F., Lee, C.G., Hurwitz, J., Parvin, J.D. and Montminy, M. (1997). RNA helicase A mediates association of CBP with RNA polymerase II. *Cell* 90(6):1107–1112.

[258] Felzien, L.K., Farrell, S., Betts, J.C., Mosavin, R. and Nabel, G.J. (1999). Specificity of cyclin E-Cdk2, TFIIB, and E1A interactions with a common domain of the p300 coactivator. *Mol Cell Biol* 19(6):4241–4246.

[259] Ogryzko, V.V., Schiltz, R.L., Russanova, V., Howard, B.H. and Nakatani, Y. (1996). The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* 87(5):953–959.

[260] Yun, J., Song, S.H., Park, J., Kim, H.P., Yoon, Y.K., Lee, K.H., Han, S.W., Oh, D.Y., Im, S.A., Bang, Y.J. et al. (2012). Gene silencing of EREG mediated by DNA methylation and histone modification in human gastric cancers. *Lab Invest* 92(7):1033–1044.

[261] Gius, D., Cui, H., Bradbury, C.M., Cook, J., Smart, D.K., Zhao, S., Young, L., Brandenburg, S.A., Hu, Y., Bisht, K.S. et al. (2004). Distinct effects on gene expression of chemical and genetic manipulation of the cancer epigenome revealed by a multimodality approach. *Cancer Cell* 6(4):361–371.

[262] Murthy, A., Defamie, V., Smookler, D.S., Grappa, M.A.D., Horiuchi, K., Federici, M., Sibilia, M., Blobel, C.P. and Khokha, R. (2010). Ectodomain shedding of EGFR ligands and TNFR1 dictates hepatocyte apoptosis during fulminant hepatitis in mice. *J Clin Invest* 120(8):2731–2744.

[263] Rauch, T.A., Wu, X., Zhong, X., Riggs, A.D. and Pfeifer, G.P. (2009). A human B cell methylome at 100-base pair resolution. *Proc Natl Acad Sci U S A* 106(3):671–678.

[264] Easwaran, H.P., Neste, L.V., Cope, L., Sen, S., Mohammad, H.P., Pageau, G.J., Lawrence, J.B., Herman, J.G., Schuebel, K.E. and Baylin, S.B. (2010). Aberrant silencing of cancer-related genes by CpG hypermethylation occurs independently of their spatial organization in the nucleus. *Cancer Res* 70(20):8015–8024.

[265] McBrian, M.A., Behbahan, I.S., Ferrari, R., Su, T., Huang, T.W., Li, K., Hong, C.S., Christofk, H.R., Vogelauer, M., Seligson, D.B. et al. (2013). Histone acetylation regulates intracellular pH. *Mol Cell* 49(2):310–321.

[266] Maitre, M., Ciesielski, L. and Mandel, P. (1974). Effect of 2-methyl 2-ethyl caproic acid and 2-2-dimethyl valeric acid on audiogenic seizures and brain gamma aminobutyric acid. *Biochem Pharmacol* 23(17):2363–2368.

[267] Eickholt, B.J., Towers, G.J., Ryves, W.J., Eikel, D., Adley, K., Ylinen, L.M.J., Chadborn, N.H., Harwood, A.J., Nau, H. and Williams, R.S.B. (2005). Effects of valproic acid derivatives on inositol trisphosphate depletion, teratogenicity, glycogen synthase kinase-3beta inhibition, and viral replication: a screening approach for new bipolar disorder drugs derived from the valproic acid core structure. *Mol Pharmacol* 67(5):1426–1433.

[268] Defossez, P.A., Kelly, K.F., Filion, G.J.P., Pérez-Torrado, R., Magdinier, F., Menoni, H., Nordgaard, C.L., Daniel, J.M. and Gilson, E. (2005). The human enhancer blocker CTC-binding factor interacts with the transcription factor Kaiso. *J Biol Chem* 280(52):43017–43023.

[269] Natoli, G. and Andrau, J.C. (2012). Noncoding transcription at enhancers: general principles and functional models. *Annu Rev Genet* 46:1–19.

[270] Kornienko, A.E., Guenzl, P.M., Barlow, D.P. and Pauler, F.M. (2013). Gene regulation by the act of long non-coding RNA transcription. *BMC Biol* 11:59.

[271] Latos, P.A., Pauler, F.M., Koerner, M.V., energin, H.B., Hudson, Q.J., Stocsits, R.R., Allhoff, W., Stricker, S.H., Klement, R.M., Warczok, K.E. et al. (2012). Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. *Science* 338(6113):1469–1472.

[272] Yang, X., Zhang, X., Mortenson, E.D., Radkevich-Brown, O., Wang, Y. and Fu, Y.X. (2013). Cetuximab-mediated tumor regression depends on innate and adaptive immune responses. *Mol Ther* 21(1):91–100.

[273] Srivastava, R.M., Lee, S.C., Filho, P.A.A., Lord, C.A., Jie, H.B., Davidson, H.C., López-Albaitero, A., Gibson, S.P., Gooding, W.E., Ferrone, S. et al. (2013). Cetuximab-activated natural killer and dendritic cells collaborate to trigger tumor antigen-specific T-cell immunity in head and neck cancer patients. *Clin Cancer Res* 19(7):1858–1872.

[274] Nicolantonio, F.D., Martini, M., Molinari, F., Sartore-Bianchi, A., Arena, S., Saletti, P., Dosso, S.D., Mazzucchelli, L., Frattini, M., Siena, S. et al. (2008). Wild-type BRAF is required for response to panitumumab or cetuximab in metastatic colorectal cancer. *J Clin Oncol* 26(35):5705–5712.

[275] Misale, S., Yaeger, R., Hobor, S., Scala, E., Janakiraman, M., Liska, D., Valtorta, E., Schiavo, R., Buscarino, M., Siravegna, G. et al. (2012). Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. *Nature* 486(7404):532–536.

[276] Oliveras-Ferraros, C., Cufí, S., Queralt, B., Vazquez-Martin, A., Martin-Castillo, B., de Llorens, R., Bosch-Barrera, J., Brunet, J. and Menendez, J.A. (2012). Cross-suppression of EGFR ligands amphiregulin and epiregulin and de-repression of FGFR3 signalling contribute to cetuximab resistance in wild-type KRAS tumour cells. *Br J Cancer* 106(8):1406–1414.

[277] Heltweg, B., Gatbonton, T., Schuler, A.D., Posakony, J., Li, H., Goehle, S., Kollipara, R., Depinho, R.A., Gu, Y., Simon, J.A. et al. (2006). Antitumor activity of a small-molecule inhibitor of human silent information regulator 2 enzymes. *Cancer Res* 66(8):4368–4377.

[278] Friend, C., Scher, W., Holland, J.G. and Sato, T. (1971). Hemoglobin synthesis in murine virus-induced leukemic cells in vitro: stimulation of erythroid differentiation by dimethyl sulfoxide. *Proc Natl Acad Sci U S A* 68(2):378–382.

[279] Levy, J., Terada, M., Rifkind, R.A. and Marks, P.A. (1975). Induction of erythroid differentiation by dimethylsulfoxide in cells infected with Friend virus: relationship to the cell cycle. *Proc Natl Acad Sci U S A* 72(1):28–32.

[280] Scher, W., Parkes, J. and Friend, C. (1977). Increased carbonic anhydrase activity in Friend erythroleukemia cells during DMSO-stimulated erythroid differentiation and its inhibition by BrdU. *Cell Differ* 6(5-6):285–296.

[281] Terada, M., Nudel, U., Fibach, E., Rifkind, R.A. and Marks, P.A. (1978). Changes in DNA associated with induction of erythroid differentiation by dimethyl sulfoxide in murine erythroleukemia cells. *Cancer Res* 38(3):835–840.

[282] Kumagai, K., Horikawa, T., Gotoh, A., Yamane, S., Yamada, H., Kobayashi, H., Hamada, Y., Suzuki, S. and Suzuki, R. (2010). Up-regulation of EGF receptor and its ligands, AREG, EREG, and HB-EGF in oral lichen planus. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod* 110(6):748–754.

[283] Leduc, V., Legault, V., Dea, D. and Poirier, J. (2011). Normalization of gene expression using SYBR green qPCR: a case for paraoxonase 1 and 2 in Alzheimer's disease brains. *J Neurosci Methods* 200(1):14–19.

8 Supplementary material

8.1 Primer and oligonucleotides

Table 23: Primers and oligos

Name	Sequence	Tm	Function
B_A_Pr_1_forw	ATGCTAGTTTggatccGGCTTGCAGCTAGAGGTTTAAG	59.71	Amplif. and Cloning of AREG promoter
H_A_Pr_3_rev	TAACCAGTGGaagcttACTCAAGAGGCAAAACCTGC	59.48	Amplif. and Cloning of AREG promoter
B_A_p150_forw	ATGCTAGTTTggatccTGGATTGGACCTCAATGACA	59.89	Amplif. and Cloning of AR intragen. site
H_A_p150_rev	TAACCAGTGGaagcttAGCCAGGTATTTGTGGTTCG	59.99	Amplif. and Cloning of AR intragen. site
B_A_p220_forw	ATGCTAGTTTggatccCAGAAGAGTATGATAACGAACCACA	59.61	Amplif. and Cloning of AR intragen. site
H_A_p220_rev	TAACCAGTCCaagcttCCACATCTCACAGGCCATATT	59.83	Amplif. and Cloning of AR intragen. site
AREG_Ampl_forw	TGGATTGGACCTCAATGACA	59.89	Amplif. and Cloning of AR intragen. site
AREG_Ampl_rev	CCACATCTCACAGGCCATATT	59.83	Amplif. and Cloning of AR intragen. site
Tata-Oligo 1	p-AGCTTACTAGTGGCTGAGCCTATAAAGCCTGCAGC	–	Cloning of Tata-BS into pCpGl-basic
Tata-Oligo 2	p-CATGGCTGCAGGCTTTATAGGCTCAGCCACTAGTA	–	Cloning of Tata-BS into pCpGl-basic
EF1-forw	ATGCTAGTTTaagcttACTAGTGGAGAAGAGCATGCTTG	60.09	Amplif. and Cloning of EF1 promoter
EF1-rev	TAACCAGTGGccatggCTGCAGGAATGTTCACAGAGACTACTGCACT	59.05	Amplif. and Cloning of EF1 promoter
AR-attB-Pr_f	ggggacaagtttgtacaaaaaagcagcctttATGAGAGCCCCGCTGCTAC	62.80	BP Clonase II-reaction
AR-attB-Pr_r	ggggaccactttgtacaagaaagctgggttTTATGCTATAGCATGTACATTTCCATT	59.35	BP Clonase II-reaction
ARsh-attB-Pr_r	ggggaccactttgtacaagaaagctgggttTTACTGGACTGTAATAACAGCAACAG	58.93	BP Clonase II-reaction
pCpGl-seq_forw	GCAAACAGCAGATTAAAAGGAA	58.57	Sequencing of pCpGl-plasmids
pCpGl-seq_rev	GGGCAGGGCCTTTCTTAAT	60.40	Sequencing of pCpGl-plasmids
CTCFtoHINDIII_1	cttccctttATAAAGCAAAAATATGGCCTGTGAGATG	65.19	Mutagenesis in pCpGl-plasmids
CTCFtoHINDIII_2	cttaaaaggcAATCATCGACAATATAGCCAGGTATTTGTG	65.79	Mutagenesis in pCpGl-plasmids
AREG_I1_forw	TCATAGAAATGACTCAATCACAAACA	59.94	Probe-gen. for Northern blot
AREG_I1_rev	CTGCAGGGGAGGGAATAAAG	60.94	Probe-gen. for Northern blot

Name	Sequence	Tm	Function
AREG_E2_forw	CCACAGTGCTGATGGATTTG	60.11	Probe-gen. for Northern blot
AREG_E2_rev	AGCCAGGTATTTGTGGTTCG	59.99	Probe-gen. for Northern blot
M13_forw	GTAAAACGACGGCCAGT	54.70	sequencing primer
M13_rev	AACAGCTATGACCATG	47.02	sequencing primer
AR_Exon2_bs_as_forw	GTTGTTGGATTGGATTTTAATGAT	58.26	strand-specific PCR: AREG_exon_2_antisense
AR_Exon2_bs_as_rev	AAAAATCTCACTCCCTAAAAACATC	58.29	strand-specific PCR: AREG_exon_2_antisense
AR_Intron1_bs_s_forw	GGGGAGGGAATAAAGGAAGA	59.35	strand-specific PCR: AREG_intron_1_sense
AR_Intron1_bs_s_rev	TCATAAAAATAACTCAATCACAAACAA	58.26	strand-specific PCR: AREG_intron_1_sense
AR_Intron1_bs_as_forw	TTTTTGATGTGGAAATTAATGTGG	59.99	strand-specific PCR: AREG_intron_1_antisense
AR_Intron1_bs_as_rev	CATTTTTACCAACAAACAAATTCAA	59.28	strand-specific PCR: AREG_intron_1_antisense
UBE_bs_sense_forw	TGTTGGAGATTATTGTGATTGTAGA	57.78	strand-specific PCR: UBE2D2_sense
UBE_bs_sense_rev	CCCAACAAATTACCCCTTCA	59.66	strand-specific PCR: UBE2D2_sense
UBE_bs_antisense_forw	TGATAGTTTTTATTAGGGTGGAGTATTTTT	60.17	strand-specific PCR: UBE2D2_antisense
UBE_bs_antisense_rev	AACCACTATAATCATAAAATATCAAAACA	57.10	strand-specific PCR: UBE2D2_antisense
KRAS_forw	AAGGCCTGCTGAAAATGACTG	59.11	DNA-contamination test in RNA
KRAS_rev	AGAATGGTCCTGCACCAGTAA	59.01	DNA-contamination test in RNA
AREG_Sybr_forw	GTGGTGCTGTCGCTCTTGATACTC	64.58	Gene expr. analysis AREG (SybRGreen)[282]
AREG_SyBR_rev	TCAAATCCATCAGCACTGTGGTC	64.54	Gene expr. analysis AREG (SybRGreen)[282]
UBE_Sybr_forw	CAGTAATGGCAGCATTTGTCTTGA	63.49	Gene expr. analysis UBE2D2-control (SybRGreen)[283]
UBE_Sybr_rev	TCATCTGGATTGGGATCACACA	63.61	Gene expr. analysis UBE2D2-control (SybRGreen)[283]

8.2 Plasmids for promoter analysis

Table 24: pCpGl-basic and -derived plasmids: CpG free plasmids were produced to analyze the methylation depending promoter activity of the AREG promoter and the AREG exon 2.

Name	Plasmid	Insert
pCpGl-basic	pCpGl-basic	none
pCpGl-CMV-EF1	pCpGl-basic	CMV-EF1
pCpGl-AREG-e2$_{150}$	pCpGl-basic	AREG exon2 CpG p150
pCpGl-AREG-e2$_{220}$	pCpGl-basic	AREG exon2 CpG p220
pCpGl-AREG-exon2	pCpGl-basic	AREG exon2
pCpGl-AREG-exon2-R	pCpGl-basic	AREG exon2 reverse
pCpGl-AREG-promoter	pCpGl-basic	AREG promoter
pCpGl-AREG-e2$_{150}$-AREG-promoter	pCpGl-AREG-promoter	AREG exon2 CpG p150
pCpGl-AREG-e2$_{220}$-AREG-promoter	pCpGl-AREG-promoter	AREG exon2 CpG p220
pCpGl-AREG-exon2-AREG-promoter	pCpGl-AREG-promoter	AREG Exon2
pCpGl-AREG-e2$_{150}$-R-AREG-promoter	pCpGl-AREG-promoter	AREG exon2 CpG p150 reverse
pCpGl-AREG-e2$_{220}$-R-AREG-promoter	pCpGl-AREG-promoter	AREG exon2 CpG p220 reverse
pCpGl-AREG-exon2-R-AREG-promoter	pCpGl-AREG-promoter	AREG exon2 reverse
pCpGl-EF1	pCpGl-basic	EF1 promoter
pCpGl-e2$_{150}$-EF1	pCpGl-EF1	AREG exon2 CpG p150
pCpGl-e2$_{220}$-EF2	pCpGl-EF1	AREG exon2 CpG p220
pCpGl-exon2-EF1	pCpGl-EF1	AREG exon2
pCpGl-e2$_{150}$-R-EF1	pCpGl-EF1	AREG exon2 CpG p150 reverse
pCpGl-e2$_{220}$-R-EF1	pCpGl-EF1	AREG exon2 CpG p220 reverse
pCpGl-exon2-R-EF1	pCpGl-EF1	AREG exon2 reverse
pCpGl-Tata	pCpGl-basic	Tata-binding motif
pCpGl-e2$_{150}$-Tata	pCpGl-Tata	AREG exon2 CpG p150
pCpGl-e2$_{220}$-Tata	pCpGl-Tata	AREG exon2 CpG p220
pCpGl-exon2-Tata	pCpGl-Tata	AREG exon2
pCpGl-e2$_{150}$-R-Tata	pCpGl-Tata	AREG exon2 CpG p150 reverse
pCpGl-e2$_{220}$-R-Tata	pCpGl-Tata	AREG exon2 CpG p220 reverse
pCpGl-exon2-R-Tata	pCpGl-Tata	AREG exon2 reverse

8.3 Sequencing results of promoter analysis plasmids

pCpGl-basic

```
GCAAACAGCAGATTAAAAGGAATTCCTGCAGGACTAGTGGATCCAGATCTTAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATG
-----------------AGGG------GCAG--CT-GTGGATC-AGATCTTA-GCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATG
                ***       ****  ** ******* ******** **********************************************************************

GCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTG-----------------------------------
GCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGT
**********************************************************************************
```

pCpGl-AREG-e2$_{150}$

```
GGGCAGGGCCTTTCTTAATATTCTTGGCATCCTCCATGGACTAAGCTTAGCCAGGTATTTGTGGTTCGTTATCATACTCTTCTGAGTAGTCAYAGWCGGCTCCCGAGGACGGTTCACTAC
-------GCCTCGC-----------------CTTCATGCACTA-GCTTAGCCAGGTATTTGTGGTTCGTTATCATACTCTTCTGAGTAGTCATAGTCGGCTCCCGAGGACGGTTCACTAC
       ****  *               ** **** **** ********************************************** ** ************************

TAGAAGGCATTTCACTCACAGGGGAAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAACCTCAAATCCATCAGCACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGG
TAGAAGGCATTTCACTCACAGGGGAAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAACCTCAAATCCATCAGCACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAAAGTAGG
************************************************************************************************************** ******

TGTCATTGAGGTCCAATCCAGGATCCACTAGTCCTGCAGGAATTCCTTTTAATCTGCTGT
TGTCATTGACGTCCAATCC-----------------CAGAGATCC---------------
******** ********                ***  ** *
```

pCpGl-AREG-e2$_{220}$

```
GCAAACAGCAGATTAAAAGGAATTCCTGCAGGACTAGTGGATCCCAGAAGAGTATGATAACGAACCACAAATACCTGGCTATATTGTCGATGATTCAGTCAGAGGTGAGTAGRGGATAAA
-----------------------CCTTGGGCAGCTGGTGGACCCCAGA-GAGTATGATAACGAACCACAAATACCTGGCTATATTGTCGATGATTCAGTCAGAGGTGAGTAGGGGATAAA
                      * **     ** ***** ****** ********************************************************************* *******

GCAAAAATATGGCCTGTGAGATGTGGAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAA
GCAAAAATATGGCCTGTGAGA-GTGGAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAA
******************** ************************************************************************************************

GGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTG-
GGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGA
************************************************************
```

pCpGl-AREG-exon2

```
GCAAACAGCAGATTAAAAGGAATTCCTGCAGGACTAGTGGATCCTGGATTGGACCTCAATGACACCTACTCTGGGAAGCGTGAACCATTTTCTGGGGACCACAGTGCTGATGGATTTGAG
-----------ATTTTCCGGAACTTGGGTGGGA--ACTGG-----GGATTGGACCTCAATGACACCTACTCTGGGAAGCGTGAACCATTTTCTGGGGACCACAGTGCTGATGGATTTGAG
          ***    **** *   *  ***  * ***     ************************************************************************

GTTACCTCAAGAAGTGAGATGTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGWCTRTGACTACTCAGAAGAGTATGATAACGAA
GTTACCTCAAGAAGTGAGATGTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGACTATGACTACTCAGAAGAGTATGATAACGAA
********************************************************************************** ** ****************************

CCACAAATACCTGGCTATATTGTCGATGATTCAGTCAGAGGTGAGTAGRGGATAAAGCAAAAATATGGCCTGTGAGATGTGGAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAA
CCACAAATACCTGGCTATATTGTCGATGATTCAGTCAGAGGTGAGTAGGGGATAAAGCAAAAATATGGCCTGTGAGATGTGGAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAA
********************************************** ************************************************************************

GGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTG-----
GGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTG
**************************************************************************************************************

```

pCpGl-AREG-exon2-R

```
-------------------------CA------------TCCACATCTCACAGGCCATATTTTTGCTTTATCCCCTACTCACCTCTGACTGAATCATCGACAATATAGCCAGGTATTTGTG
ACAGCAGATTAAAAGGAATTCCTGCAGGACTAGTGGATCCCACATCTCACAGGCCATATTTTTGCTTTATCCYCTACTCACCTCTGACTGAATCATCGACAATATAGCCAGGTATTTGTG
                        **            ********************************** ********************************************

GTTCGTTATCATACTCTTCTGAGTAGTCATAGTCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCACAGGGGAAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAA
GTTCGTTATCATACTCTTCTGAGTAGTCAYAGWCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCACAGGGGAAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAA
**************************** ** ************************************************************************************

CCTCAAATCCATCAGCACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCAATCCAGATCCAGATCTTAAGCTTAGTCCATGGAGGATGCCAAGAATAT
CCTCAAATCCATCAGCACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCAATCCAGATCCAGATCTTAAGCTTAGTCCATGGAGGATGCCAAGAATAT
**************************************************************************************************************************

TAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCC
TAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACA----------------------------
************************************************************
```

pCpGl-AREG-promoter

```
GCAAACAGCAGATTAAAAGGAATTCCTGCAGGACTAGTGGATCCGGCTTGCAGCTAGAGGTTTAAGTTCCAYTTCCTCTCAGCGAATCCTTACGCACGAGGGAGGCGGGGCGTGTGTCCT
-----------------------CCCTTCCAGGCTGGTGGA--CGGCTTGCAGCTAGAGGTTTAAGTTCCACTTCCTCTCAGCGAATCCTTACGCACGAGGGAGGCGGGGCGTGTGTCCT
                       *** *  * ** *****  *************************** ***********************************************

CCGCGCGTGGTTTTCGGGTRGCACCTTCTGGGGCGCCGCCTGCCTCCACCCACGGCCGGGCCTTGACGTCATGGGCTGCGGCCCCCTCCCGGCTGAGCCTATAAAGCGGCAGGTGCGCGC
CCGCGCGTGGTTTTCGGGTAGCACCTTCTGGGGCGCCGCCTGCCTCCACCCACGGCCGGGCCTTGACGTCATGGGCTGCGGCCCCCTCCCGGCTGAGCCTATAAAGCGGCAGGTGCGCGC
****************** ************************************************************************************************

CGCCCTACAGACGTTCGCACACCTGGGTGCCAGCGCCCCAGAGGTCCCGGGACAGCCCGAGGCGCCGCGCCCGCCGCCCCGAGCTCCCCAAGCCTTCGAGAGCGGCGCACMCTCCCGGTC
CGCCCTACAGACGTTCGCACACCTGGGTGCCAGCGCCCCAGAGGTCCCGGGACAGCCCGAGGCGCCGCGCCCGCCGCCCCGAGCTCCCCAAGCCTTCGAGAGCGGCGCACACTCCCGGTC
************************************************************************************************************ *********

TCCACTCGCTCTTCCAACACCCGCTCGTTTTGGCGGCAGCTCGTGTCCCAGAGACCGAGTTGCCCCAGAGACCGAGACGCCGCCGCYGCGAAGGACCAATGAGAGCCCCGCTGCTACCGC
TCCACTCGCTCTTCCAACACCCGCTCGTTTTGGCGGCAGCTCGTGTCCCAGAGACCGAGTTGCCCCAGAGACCGAGACGCCGCCGCTGCGAAGGACCAATGAGAGCCCCGCTGCTACCGC
************************************************************************************ ******************************

CGGCGCCRGTGGTGCTGTCGCTCTTGATACTCGGCTCAGGTGAGGATTCACCGGCGCTGAACTGCTGGGCTCTCCTCCCATGGCAGGTTTTGCCTCTTGAGTAAGCTTAGTCCATGGAGG
CGGCGCCGGTGGTGCTGTCGCTCTTGATACTCGGCTCAGGTGAGGATTCACCGGCGCTGAACTGCTGGGCTCTCCTCCCATGGCAGGTTTTGCCTCTTGAGTAAGCTTAGTCCATGGAGG
******* ************************************************************************************************************

ATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCA
ATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCA
************************************************************************************************************************

CTGATGCTCACATTG---------------------------------------------
CTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGC
**************
```

pCpGl-AREG-e2$_{150}$-AREG-promoter

```
 CCTGCAGGACTAGTGGATCGATCCTGGATTGGACCTCAATGACACCTACTCTGGGAAGCGTGAACCATTTTCTGGGGACCACAGTGCTGATGGATTTGAGGTTACCTCAAGAAGTGAGA
 ---CNNNNNNNNNNNNNNCGATCCTGGNTTGGACCTCAATGACACCTACTCTGGGAAGCGTGAACCATTTTCTGGGGACCACAGTGCTGATGGATTTGAGGTTACCTCAAGAAGTGAGA
                  ********* ********************************************************************************************

TGTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGWCTRTGACTACTCAGAAGAGTATGATAACGAACCACAAATACCTGGCTAAG
TGTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGACTATGACTACTCAGAAGAGTATGATAACGAACCACAAATACCTGGCTAAG
************************************************************** ** ******************************************************

CTGATCCGGCTTGCAGCTAGAGGTTTAAGTTCCAYTTCCTCTCAGCGAATCCTTACGCACGAGGGAGGCGGGGCGTGTGTCCTCCGCGCGCGTGGTTTTCGGGTRGCACCTTCTGGGGCGCC
CTGATCCGGCTTGCAGCTAGAGGTTTAAGTTCCACTTCCTCTCAGCGAATCCTTACGCACGAGGGAGGCGGGGCGTGTGTCCTCCGCGCGCGTGGTTTTCGGGTAGCACCTTCTGGGGCGCC
********************************** ************************************************************************ ****************

GCCTGCCTCCACCCACGGCCGGGCCTTGACGTCATGGGCTGCGGCCCCCTCCCGGCTGAGCCTATAAAGCGGCAGGTGCGCGCCGCCCTACAGACGTTCGCACACCTGGGTGCCAGCGCC
GCCTGCCTCCACCCACGGCCGGGCCTTGACGTCATGGGCTGCGGCCCCCTCCCGGCTGAGCCTATAAAGCGGCAGGTGCGCGCCGCCCTACAGACGTTCGCACACCTGGGTGCCAGCGCC
************************************************************************************************************************

CCAGAGGTCCCGGGACAGCCCGAGGCGCCGCGCCCGCCGCCCCGAGCTCCCCAAGCCTTCGAGAGCGGCGCACMCTCCCGGTCTCCACTCGCTCTTCCAACACCCGCTCGTTTTGGCGGC
CCAGAGGTCCCGGGACAGCCCGAGGCGCCGCGCCCGCCGCCCCGAGCTCCCCAAGCCTTCGAGAGCGGCGCACACTCCCGGTCTCCACTCGCTCTTCCAACACCCGCTCGTTTTGGCGGC
************************************************************** ************************************************************

AGCTCGTGTCCCAGAGACCGAGTTGCCCCAGAGACCGAGACGCCGCCGCYGCGAAGGACCAATGAGAGCCCCGCTGCTACCGCCGGCGCCRGTGGTGCTGTCGCTCTTGATACTCGGCTC
AGCTCGTGTCCCAGAGACCGAGTTGCCCCAGAGACCGAGACGCCGCCGCTGCGAAGGACCAATGAGAGCCCCGCTGCTACCGCCGGCGCCGGTGGTGCTGTCGCTCTTGATACTCGGCTC
********************************************** ************************************ ****************************

AGGTGAGGATTCACCGGCGCTGAACTGCTGGGCTCTCCTCCCATGGCAGGTTTTGCCTCTTGAGTAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTA
AGGTGAGGATTCACCGGCGCTGAACTGCTGGGCTCTCCTCCCATGGCAGGTTTTGCCTCTTGAGTAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTA
************************************************************************************************************************

CCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAA-GGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTG
CCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACNAAGGCCATGAAGAGGTATGCCCTGNNCCCTNN-ACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTG
************************************ * ********************  ****   ************************************************

AATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAG
AATACTTTGAGATGTCTGTGAGGCTGGNAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGNNGNGTGCTCTGANA-CTCTCTC-AGTNNN--ATGCCTGTGNTAGNAN
************************** ********************************************  * ********* * ******* ***     ********* *** *
```

pCpGl-AREG-e2$_{220}$-AREG-promoter

```
CCTGCAGGACTAGTGGATCGATCCCAGAAGAGTATGATAACGAACCACAAATACCTGGCTATATTGTCGATGATTCAGTCAGAGGTGAGTAGRGGATAAAGCAAAAATATGGCCTGTGAG
--CNNNNNNNNNNNGNNNCGATCCCNGAAGAGTATGATAACGAACCACAAATACCTGGCTATATTGTCGATGATTCAGTCAGAGGTGAGTAGGGGATAAAGCAAAAATATGGCCTGTGAG
             *   ******* ********************************************************************* ************************

ATGTGGAAGCTGATCCGGCTTGCAGCTAGAGGTTTAAGTTCCAYTTCCTCTCAGCGAATCCTTACGCACGAGGGAGGCGGGGCGTGTGTCCTCCGCGCGTGGTTTTCGGGTRGCACCTTC
ATGTGGAAGCTGATCCGGCTTGCAGCTAGAGGTTTAAGTTCCACTTCCTCTCAGCGAATCCTTACGCACGAGGGAGGCGGGGCGTGTGTCCTCCGCGCGTGGTTTTCGGGTAGCACCTTC
****************************************** ************************************************************************ ********

TGGGGCGCCGCCTGCCTCCACCCACGGCCGGGCCTTGACGTCATGGGCTGCGGCCCCCTCCCGGCTGAGCCTATAAAGCGGCAGGTGCGCGCCGCCCTACAGACGTTCGCACACCTGGGT
TGGGGCGCCGCCTGCCTCCACCCACGGCCGGGCCTTGACGTCATGGGCTGCGGCCCCCTCCCGGCTGAGCCTATAAAGCGGCAGGTGCGCGCCGCCCTACAGACGTTCGCACACCTGGGT
************************************************************************************************************************

GCCAGCGCCCCAGAGGTCCCGGGACAGCCCGAGGCGCCGCGCCCGCCGCCCCGAGCTCCCCAAGCCTTCGAGAGCGGCGCACMCTCCCGGTCTCCACTCGCTCTTCCAACACCCGCTCGT
GCCAGCGCCCCAGAGGTCCCGGGACAGCCCGAGGCGCCGCGCCCGCCGCCCCGAGCTCCCCAAGCCTTCGAGAGCGGCGCACACTCCCGGTCTCCACTCGCTCTTCCAACACCCGCTCGT
```

```
********************************************************************************* ***********************************

TTTGGCGGCAGCTCGTGTCCCAGAGACCGAGTTGCCCCAGAGACCGAGACGCCGCCGCYGCGAAGGACCAATGAGAGCCCCGCTGCTACCGCCGGCGCCRGTGGTGCTGTCGCTCTTGAT
TTTGGCGGCAGCTCGTGTCCCAGAGACCGAGTTGCCCCAGAGACCGAGACGCCGCCGCTGCGAAGGACCAATGAGAGCCCCGCTGCTACCGCCGGCGCCGGTGGTGCTGTCGCTCTTGAT
******************************************************** ************************************ ******************

ACTCGGCTCAGGTGAGGATTCACCGGCGCTGAACTGCTGGGCTCTCCTCCCATGGCAGGTTTTGCCTCTTGAGTAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGC
ACTCGGCTCAGGTGAGGATTCACCGGCGCTGAACTGCTGGGCTCTCCTCCCATGGCAGGTTTTGCCTCTTGAGTAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGC
*********************************************************************************************************************

CCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCAC
CCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCAC
*********************************************************************************************************************

CTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGT
CTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGT
*********************************************************************************************************************

GTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGT-GAGCAAGAAGG
GTTAGNAGCCCTGTTCATTGNNNTGGCTGTGGNCCCTGCCAATGACATCTANNATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCNN-CTGNGGNCNTTNNTGAGCANA----
***** **************   ********* ******************  ************************************   *** ** * **   *****
```

pCpGl-AREG-exon2-AREG-promoter

```
GCAAACAGCAGATTAAAAGGAATTCCTGCAGGACTAGTGGATCGATCCTGGATTGGACCTCAATGACACCTACTCTGGGAAGCGTGAACCATTTTCTGGGGACCACAGTGCTGATGGATT
------------------------NNNNNNNNNNNNNNNGNNNCGATCCTGGATTGGACCTCAATGACACCTACTCTGGGAAGCGTGAACCATTTTCTGGGGACCACAGTGCTGATGGATT
                                    *   ************************************************************************

TGAGGTTACCTCAAGAAGTGAGATGTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGWCTRTGACTACTCAGAAGAGTATGATAA
TGAGGTTACCTCAAGAAGTGAGATGTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGACTATGACTACTCAGAAGAGTATGATAA
********************************************************************************************** ** ************************

CGAACCACAAATACCTGGCTATATTGTCGATGATTCAGTCAGAGGTGAGTAGRGGATAAAGCAAAAATATGGCCTGTGAGATGTGGAAGCTGATCCGGCTTGCAGCTAGAGGTTTAAGTT
CGAACCACAAATACCTGGCTATATTGTCGATGATTCAGTCAGAGGTGAGTAGGGGATAAAGCAAAAATATGGCCTGTGAGATGTGGAAGCTGATCCGGCTTGCAGCTAGAGGTTTAAGTT
************************************************** ************************************************************************

CCAYTTCCTCTCAGCGAATCCTTACGCACGAGGGAGGCGGGGGCGTGTGTCCTCCGCGCGTGGTTTTCGGGTRGCACCTTCTGGGGCGCCGCCTGCCTCCACCCACGGCCGGGCCTTGACG
CCACTTCCTCTCAGCGAATCCTTACGCACGAGGGAGGCGGGGCGTGTGTCCTCCGCGCGTGGTTTTCGGGTAGCACCTTCTGGGGCGCCGCCTGCCTCCACCCACGGCCGGGCCTTGACG
*** ********************************************************** ************************************************

TCATGGGCTGCGGCCCCCTCCCGGCTGAGCCTATAAAGCGGCAGGTGCGCGCCGCCCTACAGACGTTCGCACACCTGGGTGCCAGCGCCCCAGAGGTCCCGGGACAGCCCGAGGCGCCGC
TCATGGGCTGCGGCCCCCTCCCGGCTGAGCCTATAAAGCGGCAGGTGCGCGCCGCCCTACAGACGTTCGCACACCTGGGTGCCAGCGCCCCAGAGGTCCCGGGACAGCCCGAGGCGCCGC
*********************************************************************************************************************

GCCCGCCGCCCCGAGCTCCCCAAGCCTTCGAGAGCGGCGCACMCTCCCGGTCTCCACTCGCTCTTCCAACACCCGCTCGTTTTGGCGGCAGCTCGTGTCCCAGAGACCGAGTTGCCCCAG
GCCCGCCGCCCCGAGCTCCCCAAGCCTTCGAGAGCGGCGCACACTCCCGGTCTCCACTCGCTCTTCCAACACCCGCTCGTTTTGGCGGCAGCTCGTGTCCCAGAGACCGAGTTGCCCCAG
****************************************** ************************************************************************

AGACCGAGACGCCGCCGCYGCGAAGGACCAATGAGAGCCCCGCTGCTACCGCCGGCGCCRGTGGTGCTGTCGCTCTTGATACTCGGCTCAGGTGAGGATTCACCGGCGCTGAACTGCTGG
AGACCGAGACGCCGCCGCTGCGAAGGACCAATGAGAGCCCCGCTGCTACCGCCGGCGCCGGTGGTGCTGTCGCTCTTGATACTCGGCTCANGTGAGGATTCACCGGCGCTGAACTGCTGG
****************** ************************************** **************************** ****************************

GCTCTCCTCCCATGGCAGGTTTTGCCTCTTGAGTAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAA
GCTCTCCTCCCATGGCAGGTTTTGCCTCTTGAGTAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAA
*********************************************************************************************************************

CTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTG-----------------------
CTGCACAAGGCCATGAAGAGGTATGCCCTGGNCCCTGGNACCATTGCCTTCACTGATGCTCACATTGAGGTGNCATCACCTATGCTGANA
******************************** ****** ***************************
```

pCpGl-AREG-e2$_{150}$-R-AREG-promoter

```
CCTGCAGGACTAGTGGATCAGCTTAGCCAGGTATTTGTGGTTCGTTATCATACTCTTCTGAGTAGTCAYAGWCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCACAGGGG
--CNNNNNNNNNNNNNNNNNGCTTAGCCNGGTATTTGTGGTTCGTTATCATACTCTTCTGAGTAGTCATAGTCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCACAGGGG
                    ******** ************************************ ** ********************************************

AAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAACCTCAAATCCATCAGCACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCAATCCAGGAT
AAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAACCTCAAATCCATCAGCACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCAATCCAGGAT
*********************************************************************************************************************

CGATCCGGCTTGCAGCTAGAGGTTTAAGTTCCAYTTCCTCTCAGCGAATCCTTACGCACGAGGGAGGCGGGGCGTGTGTCCTCCGCGCGTGGTTTTCGGGTRGCACCTTCTGGGGCGCCG
CGATCCGGCTTGCAGCTAGAGGTTTAAGTTCCACTTCCTCTCAGCGAATCCTTACGCACGAGGGAGGCGGGGCGTGTGTCCTCCGCGCGTGGTTTTCGGGTAGCACCTTCTGGGGCGCCG
********************************** ************************************************************************ ******************

CCTGCCTCCACCCACGGCCGGGCCTTGACGTCATGGGCTGCGGCCCCCTCCCGGCTGAGCCTATAAAGCGGCAGGTGCGCGCCGCCCTACAGACGTTCGCACACCTGGGTGCCAGCGCCC
CCTGCCTCCACCCACGGCCGGGCCTTGACGTCATGGGCTGCGGCCCCCTCCCGGCTGAGCCTATAAAGCGGCAGGTGCGCGCCGCCCTACAGACGTTCGCACACCTGGGTGCCAGCGCCC
*********************************************************************************************************************

CAGAGGTCCCGGGACAGCCCGAGGCGCCGCGCCCGCCGCCCCGAGCTCCCCAAGCCTTCGAGAGCGGCGCACMCTCCCGGTCTCCACTCGCTCTTCCAACACCCGCTCGTTTTGGCGGCA
CAGAGGTCCCGGGACAGCCCGAGGCGCCGCGCCCGCCGCCCCGAGCTCCCCAAGCCTTCGAGAGCGGCGCACACTCCCGGTCTCCACTCGCTCTTCCAACACCCGCTCGTTTTGGCGGCA
********************************************************************** ************************************************
```

```
GCTCGTGTCCCAGAGACCGAGTTGCCCCAGAGACCGAGACGCCGCCGCYGCGAAGGACCAATGAGAGCCCCGCTGCTACCGCCGGCGCCRGTGGTGCTGTCGCTCTTGATACTCGGCTCA
GCTCGTGTCCCAGAGACCGAGTTGCCCCAGAGACCGAGACGCCGCCGCTGCGAAGGACCAATGAGAGCCCCGCTGCTACCGCCGGCGCCGGTGGTGCTGTCGCTCTTGATACTCGGCTCA
********************************************** ************************************ ****************************

GGTGAGGATTCACCGGCGCTGAACTGCTGGGCTCTCCTCCCATGGCAGGTTTTGCCTCTTGAGTAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTAC
GGTGAGGATTCACCGGCGCTGAACTGCTGGGCTCTCCTCCCATGGCAGGTTTTGCCTCTTGAGTAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTAC
************************************************************************************************************************

CCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAA
CCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTNNCCCTNN--ACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAA
************************************************************   **     **************************************************

TACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCT
TACTTTGAGATGTCTGTGAGGCTGGCAGAN-CCATGAAAAGATATGGACTGAN----ACNNNCAGNNNTGTGGNNNGCTCTGANA-CTCT
*****************************  **********************     *    ***   *****   ******* * ****
```

pCpGl-AREG-e2$_{220}$-R-AREG-promoter

```
CCTGCAGGACTAGTGGATCAGCTTCCACATCTCACAGGCCATATTTTTGCTTTATCCYCTACTCACCTCTGACTGAATCATCGACAATATAGCCAGGTATTTGTGGTTCGTTATCATACT
---NNNNNNNNNNNNNNNNNNCTTCNCATCTCACAGGCCATATTTTTGCTTTATCCCCTACTCACCTCTGACTGAATCATCGACAATATAGCCAGGTATTTGTGGTTCGTTATCATACT
                     * * ***************************** ************************************************************

CTTCTGGGATCGATCCGGCTTGCAGCTAGAGGTTTAAGTTCCAYTTCCTCTCAGCGAATCCTTACGCACGAGGGAGGCGGGGCGTGTGTCCTCCGCGCGTGGTTTTCGGGTRGCACCTTC
CTTCTGGGATCGATCCGGCTTGCAGCTAGAGGTTTAAGTTCCACTTCCTCTCAGCGAATCCTTACGCACGAGGGAGGCGGGGCGTGTGTCCTCCGCGCGTGGTTTTCGGGTAGCACCTTC
****************************************** ************************************************************************ ********

TGGGGCGCCGCCTGCCTCCACCCACGGCCGGGCCTTGACGTCATGGGCTGCGGCCCCCTCCCGGCTGAGCCTATAAAGCGGCAGGTGCGCGCCGCCCTACAGACGTTCGCACACCTGGGT
TGGGGCGCCGCCTGCCTCCACCCACGGCCGGGCCTTGACGTCATGGGCTGCGGCCCCCTCCCGGCTGAGCCTATAAAGCGGCAGGTGCGCGCCGCCCTACAGACGTTCGCACACCTGGGT
************************************************************************************************************************

GCCAGCGCCCCAGAGGTCCCGGGACAGCCCGAGGCGCCGCGCCCGCCGCCCCGAGCTCCCCAAGCCTTCGAGAGCGGCGCACMCTCCCGGTCTCCACTCGCTCTTCCAACACCCGCTCGT
GCCAGCGCCCCAGAGGTCCCGGGACAGCCCGAGGCGCCGCGCCCGCCGCCCCGAGCTCCCCAAGCCTTCGAGAGCGGCGCACACTCCCGGTCTCCACTCGCTCTTCCAACACCCGCTCGT
************************************************************************** ************************************

TTTGGCGGCAGCTCGTGTCCCAGAGACCGAGTTGCCCCAGAGACCGAGACGCCGCCGCYGCGAAGGACCAATGAGAGCCCCGCTGCTACCGCCGGCGCCRGTGGTGCTGTCGCTCTTGAT
TTTGGCGGCAGCTCGTGTCCCAGAGACCGAGTTGCCCCAGAGACCGAGACGCCGCCGCTGCGAAGGACCAATGAGAGCCCCGCTGCTACCGCCGGCGCCGGTGGTGCTGTCGCTCTTGAT
********************************************************** ************************************ ********************

ACTCGGCTCAGGTGAGGATTCACCGGCGCTGAACTGCTGGGCTCTCCTCCCATGGCAGGTTTTGCCTCTTGAGTAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGC
ACTCGGCTCAGGTGAGGATTCACCGGCGCTGAACTGCTGGGCTCTCCTCCCATGGCAGGTTTTGCCTCTTGAGTAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGC
************************************************************************************************************************

CCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCAC
CCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGNTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGANGTGGACATCAC
****************************************************************** ********************************** ***********

CTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGA-TTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTG
CTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGAATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTG
************************************************************************ ************************************************

TGTTAGGAGCCCTGTT-CATTGGAGT-GGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAA
TGTTAGNAGCCCTGNTTCATTGGNNNTGGCTGTNGNCCCTGCCAATGACATCTACAATGANAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTNNGGTCTTTGNGAGCANGNN
****** ******* * ******    ****** * ********************** ************************************  ******** ***** *
```

pCpGl-AREG-exon2-R-AREG-promoter

```
CCTGCAGGACTAGTGGATCAGCTTCCACATCTCACAGGCCATATTTTTGCTTTATCCYCTACTCACCTCTGACTGAATCATCGACAATATAGCCAGGTATTTGTGGTTCGTTATCATACT
-NNNNNNNNNANNNNNNTNNGCTTC-ACATCTCACAGGCCATATTTTTGCTTTATCCCCTACTCACCTCTGACTGAATCATCGACAATATAGCCAGGTATTTGTGGTTCGTTATCATACT
                 *  ***** ****************************** ************************************************************

CTTCTGAGTAGTCAYAGWCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCACAGGGGAAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAACCTCAAATCCATCAG
CTTCTGAGTAGTCATAGTCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCACAGGGGAAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAACCTCAAATCCATCAG
************* ** ************************************************************************************************

CACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCAATCCAGGATCGATCCGGCTTGCAGCTAGAGGTTTAAGTTCCAYTTCCTCTCAGCGAATCCTTA
CACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCAATCCAGGATCGATCCGGCTTGCAGCTAGAGGTTTAAGTTCCACTTCCTCTCAGCGAATCCTTA
************************************************************************************************ ******************

CGCACGAGGGAGGCGGGGCGTGTGTCCTCCGCGCGTGGTTTTCGGGTRGCACCTTCTGGGGCGCCGCCTGCCTCCACCCACGGCCGGGCCTTGACGTCATGGGCTGCGGCCCCCTCCGG
CGCACGAGGGAGGCGGGGCGTGTGTCCTCCGCGCGTGGTTTTCGGGTAGCACCTTCTGGGGCGCCGCCTGCCTCCACCCACGGCCGGGCCTTGACGTCATGGGCTGCGGCCCCCTCCGG
********************************************** ************************************************************************

CTGAGCCTATAAAGCGGCAGGTGCGCGCCGCCCTACAGACGTTCGCACACCTGGGTGCCAGCGCCCCAGAGGTCCCGGGACAGCCCGAGGCGCCGCCCGCCGCCCCGAGCTCCCCAAG
CTGAGCCTATAAAGCGGCAGGTGCGCGCCGCCCTACAGACGTTCGCACACCTGGGTGCCAGCGCCCCAGAGGTCCCGGGACAGCCCGAGGCGCCGCCCGCCGCCCCGAGCTCCCCAAG
************************************************************************************************************************

CCTTCGAGAGCGGCGCACMCTCCCGGTCTCCACTCGCTCTTCCAACACCCGCTCGTTTTGGCGGCAGCTCGTGTCCCAGAGACCGAGTTGCCCCAGAGACCGAGACGCCGCCGCYGCGAA
CCTTCGAGAGCGGCGCACACTCCCGGTCTCCACTCGCTCTTCCAACACCCGCTCGTTTTGGCGGCAGCTCGTGTCCCAGAGACCGAGTTGCCCCAGAGACCGAGACGCCGCCGCTGCGAA
****************** ************************************************************************************** *****

GGACCAATGAGAGCCCCGCTGCTACCGCCGGCGCCRGTGGTGCTGTCGCTCTTGATACTCGGCTCAGGTGAGGATTCACCGGCGCTGAACTGCTGGGCTCTCCTCCCATGGCAGGTTTTG
GGACCAATGAGAGCCCCGCTGCTACCGCCGGCGCCGGTGGTGCTGTCGCTCTTGATACTCGGCTCANGTGAGGATTCACCGGCGCTGAACTGCTGGGCTCTCCTCCCATGGCAGGTTTTG
```

```
********************************** *************************** ************************************************

CCTCTTGAGTAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTAT
CCTCTTGAGTAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAANATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTAT
************************************************************************ ****************************************

GCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAAC
GCCCTNNCCCTN--NACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAN-ACTTTGAGATGTCTGTGAG-CTGGNANAN-CCATGAAAAGANATGGNCTGAAC
*****   **     ************************************************  ****************** **** * *  *********** **** ******
```

pCpGl-EF1

```
GCAAACAGCAGATTAAAAGGAATTCCTGCAGGACTAGTGGATCCAGATCTTAAGCTTACTAGTGGAGAAGAGCATGCTTGAGGGCTGAGTGCCCCTCAGTGGGCAGAGAGCACATGGCCC
-------------------NNNNNNNNNNNNNNNNNNNNNNNNNGATCTTAAGCTTAC-AGTGGAGAAGAGCATGCTTGAGGGCTGAGTGCCCCTCAGTGGGCAGAGAGCACATGGCCC
                                     ************* ************************************************************

ACAGTCCCTGAGAAGTTGGGGGGAGGGGTGGGCAATTGAACTGGTGCCTAGAGAAGGTGGGGCTTGGGTAAACTGGGAAAGTGATGTGGTGTACTGGCTCCACCTTTTTCCCCAGGGTGG
ACAGTCCCTGAGAAGTTGGGGGGAGGGGTGGGCAATTGAACTGGTGCCTAGAGAAGGTGGGGCTTGGGTAAACTGGGAAAGTGATGTGGTGTACTGGCTCCACCTTTTTCCCCAGGGTGG
************************************************************************************************************************

GGGAGAACCATATATAAGTGCAGTAGTCTCTGTGAACATTCCTGCAGCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGA
GGGAGAACCATATATAAGTGCAGTAGTCTCTGTGAACATTCCTGCAGCCATG-AGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGA
************************************************** ************************************************************************

GCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGC
GCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGC
************************************************************************************************************************

AGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCCCC
AGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCCCC
************************************************************************************************************************

TGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTGCC
TGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTGCC
************************************************************************************************************************

CATCATCCAGAAGATCATCATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGACTTTGTGCCTGA
CATCATCCAGAAGATCATCATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGACTTTGTGCCTGA
************************************************************************************************************************

GAGCTTTGACAGGGACAAGACCATTGCTCTGATTATGAACAGCTCTGGCTCCACTGGACTGCCCAAAGGTGTGGCTCTGCCCCACAGAACTGCTTG-TGTGAGATTCAGCCATGCCAGAG
GAGCTTTGACAGGGACAAGACCATTGCTCTGATTATGAACAGCTCTGGCTCCACTGGACTGCCCAAAGGTGTGGCTCTGCCCCACAGAACTGCTTNGTGTGAGATTCAGCCATGCCAGAG
************************************************************************************************  **********************

ACCCCATCTTTGGCAACCAGATCATCCCTGACACTGCCATCCTGTCTGTGGTTCCATTCCATCATGGCTTTGGCATGTTCACAACACTGGGGTACCTGATCTGTGGCTTCAGAGTGGTGC
ACCCCATCTTTGGCAACCAGATCATCCCTGACACTGCCATCCTGTCTGTGGTTCCATTCCATCATGGCTTTGGCATGTTCACAACACTGGGGNACCTGATCTGNGGCTTCAGAGTGN-GC
************************************************************************************************ ********** ***********  **

TGATGTATAGGT-----------------------------------------------
TGATGTATNNGTTTGAGGNNAGCTGTTTTCTGAGNAGCCNNNNAAGACTACANATCCAGT
********  **
```

pCpGl-AREG-e2$_{150}$-EF1

```
--CCTGCAGGACTAGTGGATC------GATCCTGGATTGGACCTCAATGACACCTACTCTGGGAAGCGTGAACCATTTTCTGGGGACCACAGTGCTGATGGATTTGAGGTTACCTCAAGA
CNNNNNNNNNNNNNNNNGATCTTAGCTGATCCTGGATTGGACCTCAATGACACCTACTCTGGGAAGCGTGAACCATTTTCTGGGGACCACAGTGCTGATGGATTTGAGGTTACCTCAAGA
                 ****      ************************************************************************************************

AGTGAGATGTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGWCTRTGACTACTCAGAAGAGTATGATAACGAACCACAAATACCT
AGTGAGATGTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGACTATGACTACTCAGAAGAGTATGATAACGAACCACAAATACCT
************************************************************************** ** ************************************

GGCTAAGCTGATCCAGATCTTAAGCTTACTAGTGGAGAAGAGCATGCTTGAGGGCTGAGTGCCCCTCAGTGGGCAGAGAGCACATGGCCCACAGTCCCTGAGAAGTTGGGGGGAGGGGTG
GGCTAAGCT-------------AGCTTAC-AGTGGAGAAGAGCATGCTTGAGGGCTGAGTGCCCCTCAGTGGGCAGAGAGCACATGGCCCACAGTCCCTGAGAAGTTGGGGGGAGGGGTG
*********             ******* ************************************************************************************************

GGCAATTGAACTGGTGCCTAGAGAAGGTGGGGCTTGGGTAAACTGGGAAAGTGATGTGGTGTACTGGCTCCACCTTTTTCCCCAGGGTGGGGGAGAACCATATATAAGTGCAGTAGTCTC
GGCAATTGAACTGGTGCCTAGAGAAGGTGGGGCTTGGGTAAACTGGGAAAGTGATGTGGTGTACTGGCTCCACCTTTTTCCCCAGGGTGGGGGAGAACCATATATAAGTGCAGTAGTCTC
************************************************************************************************************************

TGTGAACATTCCTGCAGCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGC
TGTGAACATTCCTGCAGCCATG-AGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGC
********************** ************************************************************************************************

CCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACAC
CCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACAC
************************************************************************************************************************

CAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCT
CAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCT
************************************************************************************************************************
```

```
CCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAG
CCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAANCTGCCCATCATCCAGAANATCATCATCATGGNCAG
********************************************************************************** ********** ********** ***

CAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGACTTTGTGCCTGAGAGCTTTGACAGGGACAAGAC---------
CAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTNNNT-CAATGAGTATGACTTTGTGCCTGANANCTTTGACAGGGNNAGANCATTGCTCTG
************************************************************   * ********** * ********** *   *
```

pCpGl-AREG-e2$_{220}$-EF1

```
---CCTGCAGGACTAGTGGATC------GATCCCAGAAGAGTATGATAACGAACCACAAATACCTGGCTATATTGTCGATGATTCAGTCAGAGGTGAGTAGRGGATAAAGCAAAAATATG
NNNNNNNNNNNNNNGNNNNGATCTTANCTGATCCCAGAAGAGTATGATAACGAACCACAAATACCTGGCTATATTGTCGATGATTCAGTCAGAGGTGAGTAGGGGATAAAGCAAAAATATG
                ****      ********************************************************************** ******************

GCCTGTGAGATGTGGAAGCTGATCCAGATCTTAAGCTTACTAGTGGAGAAGAGCATGCTTGAGGGCTGAGTGCCCCTCAGTGGGCAGAGAGCACATGGCCCACAGTCCCTGAGAAGTTGG
GCCTGTGAGATGTGGAAGCT-------------AGCTTAC-AGTGGAGAAGAGCATGCTTGAGGGCTGAGTGCCCCTCAGTGGGCAGAGAGCACATGGCCCACAGTCCCTGAGAAGTTGG
******************             ******* ********************************************************************************

GGGGAGGGGTGGGCAATTGAACTGGTGCCTAGAGAAGGTGGGGCTTGGGTAAACTGGGAAAGTGATGTGGTGTACTGGCTCCACCTTTTTCCCCAGGGTGGGGGAGAACCATATATAAGT
GGGGAGGGGTGGGCAATTGAACTGGTGCCTAGAGAAGGTGGGGCTTGGGTAAACTGGGAAAGTGATGTGGTGTACTGGCTCCACCTTTTTCCCCAGGGTGGGGGAGAACCATATATAAGT
********************************************************************************************************************

GCAGTAGTCTCTGTGAACATTCCTGCAGCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATG
GCAGTAGTCTCTGTGAACATTCCTGCAGCCATG-AGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATG
********************************** ********************************************************************************************

AAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATAT
AAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATAT
********************************************************************************************************************

GGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAAT
GGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAAT
********************************************************************************************************************

GAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATC
GAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATC
********************************************************************************************************************

ATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGACTTTGTGCCTGAGAGCTTTGACAGGGACAAG
ATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGACTTTGTGCCTGANAGCTTTGACAGGGACAAG
********************************************************************************************************** ******************

ACCATTGCTCTGATTATGAACAGCTCTGGCTCCACTGGACTGCCCAAAGGTG--------
ACCATTGCTCTGATTATGAACAGCTCTGGCTCCACTGGNCTGCCCAAAGGTGTGGCTCTG
************************************ *************
```

pCpGl-AREG-exon2-EF1

```
CCTGCAGGACTAGTGGATCGATCCTGGATTGGACCTCAATGACACCTACTCTGGGAAGCGTGAACCATTTTCTGGGGACCACAGTGCTGATGGATTTGAGGTTACCTCAAGAAGTGAGAT
---------------NNNCGATCCTGGANTGGACCTCAATGACACCTACTCTGGGAAGCGTGAACCATTTTCTGGGGACCACAGTGCTGATGGATTTGAGGTTACCTCAAGAAGTGAGAT
                ********** ********************************************************************************************

GTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGWCTRTGACTACTCAGAAGAGTATGATAACGAACCACAAATACCTGGCTATAT
GTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGACTATGACTACTCAGAAGAGTATGATAACGAACCACAAATACCTGGCTATAT
************************************************************** ** ******************************************

TGTCGATGATTCAGTCAGAGGTGAGTAGRGGATAAAGCAAAAATATGGCCTGTGAGATGTGGAAGCTGATCCAGATCTTAAGCTTACTAGTGGAGAAGAGCATGCTTGAGGGCTGAGTGC
TGTCGATGATTCAGTCAGAGGTGAGTAGGGGATAAAGCAAAAATATGGCCTGTGAGATGTGGAAGCTGATCCAGATCTTAAGCTTAC-AGTGGAGAAGAGCATGCTTGAGGGCTGAGTGC
**************************** ************************************************** ******************************

CCCTCAGTGGGCAGAGAGCACATGGCCCACAGTCCCTGAGAAGTTGGGGGGAGGGGTGGGCAATTGAACTGGTGCCTAGAGAAGGTGGGGCTTGGGTAAACTGGGAAAGTGATGTGGTGT
CCCTCAGTGGGCAGAGAGCACATGGCCCACAGTCCCTGAGAAGTTGGGGGGAGGGGTGGGCAATTGAACTGGTGCCTAGAGAAGGTGGGGCTTGGGTAAACTGGGAAAGTGATGTGGTGT
********************************************************************************************************************

ACTGGCTCCACCTTTTTCCCCAGGGTGGGGGAGAACCATATATAAGTGCAGTAGTCTCTGTGAACATTCCTGCAGCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCT
ACTGGCTCCACCTTTTTCCCCAGGGTGGGGGAGAACCATATATAAGTGCAGTAGTCTCTGTGAACATTCCTGCAGCCATG-AGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCT
************************************************************************** **********************************

ACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTG
ACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTG
********************************************************************************************************************

AATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAG
AATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAG
********************************************************************************************************************

CCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAA
CCCTGTTCATTGNAGTGGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAANNGCCNT----
************ ********************************************************************************************  ***
```

pCpGl-AREG-e2$_{150}$-R-EF1

```
-CCTGCAGGACTAGTGGAT--------CAGCTTAGCCAGGTATTTGTGGTTCGTTATCATACTCTTCTGAGTAGTCAYAGWCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCAC
NNNNNNNNNANNNGNNNNNATCTTNNCTAGCTTAGCCAGGTATTTGTGGTTCGTTATCATACTCTTCTGAGTAGTCATAGTCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCAC
        *   *             ********************************************** ** ************************************

TCACAGGGGAAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAACCTCAAATCCATCAGCACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCA
TCACAGGGGAAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAACCTCAAATCCATCAGCACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCA
************************************************************************************************************************

ATCCAGGATCGATCCAGATCTTAAGCTTACTAGTGGAGAAGAGCATGCTTGAGGGCTGAGTGCCCCTCAGTGGGCAGAGAGCACATGGCCCACAGTCCCTGAGAAGTTGGGGGGGAGGGGT
ATCCAGGA---------------AGCTTAC-AGTGGAGAAGAGCATGCTTGAGGGCTGAGTGCCCCTCAGTGGGCAGAGAGCACATGGCCCACAGTCCCTGAGAAGTTGGGGGGGAGGGGT
********              ******* ************************************************************************************

GGGCAATTGAACTGGTGCCTAGAGAAGGTGGGGGCTTGGGTAAACTGGGAAAGTGATGTGGTGTACTGGCTCCACCTTTTTCCCCAGGGTGGGGGAGAACCATATATAAGTGCAGTAGTCT
GGGCAATTGAACTGGTGCCTAGAGAAGGTGGGGGCTTGGGTAAACTGGGAAAGTGATGTGGTGTACTGGCTCCACCTTTTTCCCCAGGGTGGGGGAGAACCATATATAAGTGCAGTAGTCT
************************************************************************************************************************

CTGTGAACATTCCTGCAGCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATG
CTGTGAACATTCCTGCAGCCATG-AGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATG
********************** ************************************************************************************************

CCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACA
CCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACA
************************************************************************************************************************

CCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGC
CCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGC
************************************************************************************************************************

TCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACA
TCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTGCCCATCATCCANAAGATCATCATCATGGNCA
************************************************************************************************ *************** **

GCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGACTTTGTGCCTGAGAGCTTTGACAGGGACAAGACCATTGCTC
GCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTNNNT-CNNTGAGTATGACTTTGNGCCTGANAGCTT-GACAGGGANA-GACCATTGCTC
************************************************************   * *  ************** ****** ***** ******** * ***********

TGAT-------------------------------TATGAACAGCTCTGGCT---------
TGATATNANAGCTCTGNNCNCTGACTGCCNANGGTNNGNNNNNCCCNNNNNACNGCTNNN
****                              * *    * *
```

pCpGl-AREG-e2$_{220}$-R-EF1

```
----CCTGCAGGACTAGTGGATC------AGCTTCCACATCTCACAGGCCATATTTTTGCTTTATCCYCTACTCACCTCTGACTGAATCATCGACAATATAGCCAGGTATTTGTGGTTCG
NNNNNNNNNNNNNNNNNNNGATCTTAGCTAGCTTCCACATCTCACAGGCCATATTTTTGCTTTATCCCCTACTCACCTCTGACTGAATCATCGACAATATAGCCAGGTATTTGTGGTTCG
                   ****      ************************************ ************************************************

TTATCATACTCTTCTGGGATCGATCCAGATCTTAAGCTTACTAGTGGAGAAGAGCATGCTTGAGGGCTGAGTGCCCCTCAGTGGGCAGAGAGCACATGGCCCACAGTCCCTGAGAAGTTG
TTATCATACTCTTCTGGGATC-------------AGCTTAC-AGTGGAGAAGAGCATGCTTGAGGGCTGAGTGCCCCTCAGTGGGCAGAGAGCACATGGCCCACAGTCCCTGAGAAGTTG
*********************             ******* ************************************************************************

GGGGGAGGGGTGGGCAATTGAACTGGTGCCTAGAGAAGGTGGGGGCTTGGGTAAACTGGGAAAGTGATGTGGTGTACTGGCTCCACCTTTTTCCCCAGGGTGGGGGAGAACCATATATAAG
GGGGGAGGGGTGGGCAATTGAACTGGTGCCTAGAGAAGGTGGGGGCTTGGGTAAACTGGGAAAGTGATGTGGTGTACTGGCTCCACCTTTTTCCCCAGGGTGGGGGAGAACCATATATAAG
************************************************************************************************************************

TGCAGTAGTCTCTGTGAACATTCCTGCAGCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCAT
TGCAGTAGTCTCTGTGAACATTCCTGCAGCCATG-AGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCAT
********************************** ************************************************************************************

GAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATA
GAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATA
************************************************************************************************************************

TGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAA
TGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAA
************************************************************************************************************************

TGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCAT
TGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCAT
************************************************************************************************************************

CATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGACTTTGTGCCTGAGAGCTTTGACAGGGACAA
CATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTNN-TTCAATGAGTATGACTTTGTGCCTGANAGCTTTGACAGGGANAA
********************************************************************   ************************* ************** **

GACCATTGCTCTGATTATGAACAGCTCTGGCTCCACTGGACTGCCCAAAGGTG-------
GACCATTGCTCTGATTATGAACAGCTCTGGCTCCACTGGNCTGCCCNAAGGTGTGGCTCT
************************************ ****** ******
```

pCpGl-AREG-exon2-R-EF1

```
CCTGCAGGACTAGTGGATCAGCTTCCACATCTCACAGGCCATATTTTTGCTTTATCCYCTACTCACCTCTGACTGAATCATCGACAATATAGCCAGGTATTTGTGGTTCGTTATCATACT
--CNNNNNNNNNNNNNNTNNGCTTCN-CA-CTCNCAGGCCATATTTTTGCTTTATCCCCTACTCACCTCTGACTGAATCATCGACAATATAGCCAGGTATTTGTGGTTCGTTATCATACT
```

```
              *  *****  ** *** ********************** ************************************************************************

CTTCTGAGTAGTCAYAGWCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCACAGGGGAAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAACCTCAAATCCATCAG
CTTCTGAGTAGTCATAGTCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCACAGGGGAAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAACCTCAAATCCATCAG
************* ** ********************************************************************************************************

CACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCAATCCAGGATCGATCCAGATCTTAAGCTTACTAGTGGAGAAGAGCATGCTTGAGGGCTGAGTGC
CACTGTGGTCCCCAGAAAGTGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCAATCCAGGATCGATCCAGATCTTAAGCTTAC-AGTGGAGAAGAGCATGCTTGAGGGCTGAGTGC
***************** ********************************************************************* ******************************

CCCTCAGTGGGCAGAGAGCACATGGCCCACAGTCCCTGAGAAGTTGGGGGGAGGGGTGGGCAATTGAACTGGTGCCTAGAGAAGGTGGGGCTTGGGTAAACTGGGAAAGTGATGTGGTGT
CCCTCAGTGGGCAGAGAGCACATGGCCCACAGTCCCTGAGAAGTTGGGGGGAGGGGTGGGCAATTGAACTGGTGCCTAGAGAAGGTGGGGCTTGGGTAAACTGGGAAAGTGATGTGGTGT
************************************************************************************************************************

ACTGGCTCCACCTTTTTCCCCAGGGTGGGGGAGAACCATATATAAGTGCAGTAGTCTCTGTGAACATTCCTGCAGCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCT
ACTGGCTCCACCTTTTTCCCCAGGGTGGGGGAGAACCATATATAAGTGCAGTAGTCTCTGTGAACATTCCTGCAGCCATG-AGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCT
************************************************************************** ************************************

ACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTG
ACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTG
************************************************************************************************************************

AATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAG
AATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTANGAG
************************************************************************************************************************ ***

CCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAA
CCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAANGGCCTGCAAA
************************************************************************************************************************ **********

AGATCCTGAATGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTG
AGATCCTGAATGTGCAGAANAANCTGCCCATCATCCAGAANATCATCATCATGGNNAGCAAGACTGACTACCAGGGCTTCCAGAGCATGNNNACCTTTGTGACCAGCCACTTACCCCCNN
****************** ** ***************** *************  *******************************   ************************

GCTTCAATGAGTATGACTTTGTGCCTGAG------AGCTTTGACAGGGACA------AGA
NNNN--ANGANTATGANTTTGNNCNGNNNCTTNGANNNNNNNACATGNNNNGATATGANN
      * ** ***** ****  *                 *** *         *
```

pCpGl-Tata

```
GCAAACAGCAGATTAAAAGGAATTCCTGCAGGACTAGTGGATCCAGATCTTAAGCTTACTAGTGGCTGAGCCTATAAAGCCTGCAGCCATGGAGGATGCCAAGAATATTAAGAAAGGCCC
--------------------NNNNNNNNNNNNNNNNNNNNNNNNNNNNATCTTAAGCTTACTAGTGGCTGAGCCTATAAAGCCTGCAGCCATGGAGGATGCCAAGAATATTAAGAAAGGCCC
                                          ************************************************************************

TGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACAT
TGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACAT
************************************************************************************************************************

CACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCC
CACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCC
************************************************************************************************************************

TGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAA
TGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAA
************************************************************************************************************************

GGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCA
GGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCA
************************************************************************************************************************

CTTACCCCCTGGCTTCAATGAGTATGACTTTGTGCCTGAGAGCTTTGACAGGGACAAGACCATTGCTCTGATTATGAACAGCTCTGGCTCCACTGGACTGCCCAAAGGTGTGGCTCTGCC
CTTACCCCCTGGCTTCAATGAGTATGACTTTGTGCCTGAGAGCTTTGACAGGGACAAGACCATTGCTCTGATTATGAACAGCTCTGGCTCCACTGGACTGCCCAAAGGTGTGGCTCTGCC
************************************************************************************************************************

CCACAGAACTGCTTGTGTGAGATTCAGCCATGCCAGAGACCCCATCTTTGGCAACCAGATCATCCCTGACACTGCCATCCTGTCTGTGGTTCCATTCCATCATGGCTTTGGCATGTTCAC
CCACAGAACTGCTTGTGTGAGATTCAGCCATGCCAGAGACCCCATCTTTGGCAACCAGATCATCCCTGACACTGCCATCCTGTCTGTGGTTCCATTCCATCATGGCTTTGGCATGTTCAC
************************************************************************************************************************

AACACTGGGGTACCTGATCTGTGGCTTCAGAGTGGTGCTGATGTATAGGT----------
AACACTGGGGTACCTGATCTGTGGCTTCAGAGTGGTGCTGATGTATAGGTTTGAGGAGGA
**************************************************
```

pCpGl-AREG-e2$_{150}$-Tata

```
CCTGCAGGACTAGTGGATCGATCCTGGATTGGACCTCAATGACACCTACTCTGGGAAGCGTGAACCATTTTCTGGGGACCACAGTGCTGATGGATTTGAGGTTACCTCAAGAAGTGAGAT
----CNNNNNNNNNNNNNCGATCCTGGATTGGACCTCAATGACACCTACTCTGGGAGGCGTGAACCATTTTCTGGGGACCACAGTGCTGATGGATTTGAGGTTACCTCAAGAAGTGAGAT
    *            ************************************ ************************************************************

GTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGWCTRTGACTACTCAGAAGAGTATGATAACGAACCACAAATACCTGGCTAAGC
GTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGACTATGACTACTCAGAAGAGTATGATAACGAACCACAAATACCTGGCTAAGC
************************************************************** ** ******************************************************
```

```
TGATCCAGATCTTAAAGCTTACTAGTGGCTGAGCCTATAAAGCCTGCAGCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGT
TGATCCAGATCTTAA-GCTTACTAGTGGCTGAGCCTATAAAGCCTGCAGCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGT
********** ************************************************************************************************************

GAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTG
GAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTG
************************************************************************************************************************

GCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCC
GCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCC
************************************************************************************************************************

CCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTG
CCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTG
************************************************************************************************************************

CCCATCATCCAGAAGATCATCATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGACTTTGTGCCT
CCCATCATCCAGAAGATCATCATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGACTTTGTGCCT
************************************************************************************************************************

GAGAGCTTTGACAGGGACAAGACCATTGCTCTGATTATGAACAGCTCTGGCTCCACTGGACTGCCCAAAGGTG----------------------------------------------
GAGAGCTTTGACAGGGACAAGACCATTGCTCTGATTATGAACAGCTCTGGCTCCACTGGACTGCCCAAAGGTGTGGCTCTGCCCCACAGAACTGCTTGTGTGAGATTCAGCCATGCCAGA
************************************************************************
```

pCpGl-AREG-e2$_{220}$-Tata

```
CCTGCAGGACTAGTGGATCGATCCCAGAAGAGTATGATAACGAACCACAAATACCTGGCTATATTGTCGATGATTCAGTCAGAGGTGAGTAGRGGATAAAGCAAAAATATGGCCTGTGAG
--NNNNNNNNNNNNNNNNCGATCCCNGAAGAGTATGATAACGAACCACAAATACCTGGCTATATTGTCGATGATTCAGTCAGAGGTGAGTAGGGGATAAAGCAAAAATATGGCCTGTGAG
                 ******* ************************************************************ **************************

ATGTGGAAGCTGATCCAGATCTTAAAGCTTACTAGTGGCTGAGCCTATAAAGCCTGCAGCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGG
ATGTGGAAGCTGATCCAGATCTTAA-GCTTACTAGTGGCTGAGCCTATAAAGCCTGCAGCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGG
************************ ************************************************************************************************

CACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTC
CACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTC
************************************************************************************************************************

TGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGT
TGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGT
************************************************************************************************************************

GGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCA
GGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCA
************************************************************************************************************************

GAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGA
GAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGA
************************************************************************************************************************

CTTTGTGCCTGAGAGCTTTGACAGGGACAAGACCATTGCTCTGATTATGAACAGCTCTGGCTCCACTGGACTGCCCAAAGGTG----------------------------------
CTTTGTGCCTGAGAGCTTTGACAGGNACAAGACCATTGCTCTGATTATGAACAGCTCTGGCTCCACTGGACTGCCCAAAGGTGTGGCTCTGCCCCACAGAACTGCTTGTGTGAGATTCAG
********************** ************************************************

**pCpGl-AREG-exon2-Tata**

CCTGCAGGACTAGTGGATCGATCCTGGATTGGACCTCAATGACACCTACTCTGGGAAGCGTGAACCATTTTCTGGGGACCACAGTGCTGATGGATTTGAGGTTACCTCAAGAAGTGAGAT
--CNNNNNNNNNNNNNNNCGATCCTGGNTTGGACCTCAATGACACCTACTCTGGGAAGCGTGAACCATTTTCTGGGGACCACAGTGCTGATGGATTTGAGGTTACCTCAAGAAGTGAGAT
                 ********* ************************************************************************************************

GTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGWCTRTGACTACTCAGAAGAGTATGATAACGAACCACAAATACCTGGCTATAT
GTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGACTATGACTACTCAGAAGAGTATGATAACGAACCACAAATACCTGGCTATAT
************************************************************** ** ******************************************

TGTCGATGATTCAGTCAGAGGTGAGTAGRGGATAAAGCAAAAATATGGCCTGTGAGATGTGGAAGCTGATCCAGATCTTAAAGCTTACTAGTGGCTGAGCCTATAAAGCCTGCAGCCATG
TGTCGATGATTCAGTCAGAGGTGAGTAGGGGATAAAGCAAAAATATGGCCTGTGAGATGTGGAAGCTGATCCAGATCTTAA-GCTTACTAGTGGCTGAGCCTATAAAGCCTGCAGCCATG
**************************** ************************************************** ************************************

GAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCC
GAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCC
************************************************************************************************************************

TTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGC
TTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGC
************************************************************************************************************************

TCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGC
TCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGC
************************************************************************************************************************

CAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAAGACTGACTACCAGGGCTTC
CAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAAGACTGACTACCAGGGCTTC
```

```
**********************************************************************************************************************

CAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGACTTTGTGCCTGAGAGCTTTGACAGGGACAAGACCATTGCTCTGATTATGAACAGCTCTGGCTCC
CAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGACTTTGTGCCTGAGAGCTTTGACAGGGACAAGACCATTGCTCTGATTATGAACAGCTCTGGCTCC
**********************************************************************************************************************

ACTGGACTGCCCAAAGGTG------------------------------------------
ACTGGNCTGCCCAAAGGTGTGGCTCTGCCCCACAGAACTGCTTGTGTGANATTCAGCCAT
***** ************
```

pCpGl-AREG-e2$_{150}$-R-Tata

```
CCTGCAGGACTAGTGGATCAGCTTAGCCAGGTATTTGTGGTTCGTTATCATACTCTTCTGAGTAGTCAYAGWCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCACAGGGG
-CNNNNNNNNANNNNNNNNNGCTTAGCCNGGTATTTGTGGTTCGTTATCATACTCTTCTGAGTAGTCATAGTCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCACAGGGG
 *                 ******** ************************************ ** ********************************************

AAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAACCTCAAATCCATCAGCACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCAATCCAGGAT
AAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAACCTCAAATCCATCAGCACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCAATCCAGGAT
**********************************************************************************************************************

CGATCCAGATCTTAAAGCTTACTAGTGGCTGAGCCTATAAAGCCTGCAGCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGT
CGATCCAGATCTTAA-GCTTACTAGTGGCTGAGCCTATAAAGCCTGCAGCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGT
*************** **********************************************************************************************************

GAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTG
GAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTG
**********************************************************************************************************************

GCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCC
GCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCC
**********************************************************************************************************************

CCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTG
CCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTG
**********************************************************************************************************************

CCCATCATCCAGAAGATCATCATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGACTTTGTGCCT
CCCATCATCCAGAAGATCATCATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGACTTTGTGCCT
**********************************************************************************************************************

GAGAGCTTTGACAGGGACAAGACCATTGCTCTGATTATGAACAGCTCTGGCTCCACTGGACTGCCCAAAGGTG----------------------------------------------
GAGAGCTTTGACAGGGACAAGACCATTGCTCTGATTATGAACAGCTCTGGCTCCACTGGACTGCCCAAAGGTGTGGCTCTGCCCCACAGAACTGCTTGTGTGANATTCAGCCATGCCAGA
**********************************************************************
```

pCpGl-AREG-e2$_{220}$-R-Tata

```
CCTGCAGGACTAGTGGATCAGCTTCCACATCTCACAGGCCATATTTTTGCTTTATCCYCTACTCACCTCTGACTGAATCATCGACAATATAGCCAGGTATTTGTGGTTCGTTATCATACT
--CNNNNNNNNNNNNNNNNAGCTTCN-CATCTCACAGGCCATATTTTTGCTTTATCCCCTACTCACCTCTGACTGAATCATCGACAATATAGCCAGGTATTTGTGGTTCGTTATCATACT
                  ******  ******************************** **********************************************************

CTTCTGGGATCGATCCAGATCTTAAAGCTTACTAGTGGCTGAGCCTATAAAGCCTGCAGCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGG
CTTCTGGGATCGATCCAGATCTTAA-GCTTACTAGTGGCTGAGCCTATAAAGCCTGCAGCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGG
********************** **********************************************************************************************

CACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTC
CACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCCTTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTC
**********************************************************************************************************************

TGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGT
TGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGCTCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGT
**********************************************************************************************************************

GGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCA
GGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGCCAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCA
**********************************************************************************************************************

GAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGA
GAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAAGACTGACTACCAGGGCTTCCAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGA
**********************************************************************************************************************

CTTTGTGCCTGAGAGCTTTGACAGGGACAAGACCATTGCTCTGATTATGAACAGCTCTGGCTCCACTGGACTGCCCAAAGGTG-----------------------------------
CTTTGTGCCTGAGAGCTTTGACAGGGACAAGACCATTGCTCTGATTATGAACAGCTCTGGCTCCACTGGACTGCCCAAAGGTGTGGCTCTGCCCCACAGAACTGCTTGTGTGAGATTCAG
**************************************************************************
```

pCpGl-AREG-exon2-R-Tata

```
CCTGCAGGACTAGTGGATCAGCTTCCACATCTCACAGGCCATATTTTTGCTTTATCCYCTACTCACCTCTGACTGAATCATCGACAATATAGCCAGGTATTTGTGGTTCGTTATCATACT
--CNNNNNNNNNNNNGNNNNAGCTTC-ACATCTCACAGGCCATATTTTTGCTTTATCCCCTACTCACCTCTGACTGAATCATCGACAATATAGCCAGGTATTTGTGGTTCGTTATCATACT
             *    ****** ******************************** **********************************************************

CTTCTGAGTAGTCAYAGWCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCACAGGGGAAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAACCTCAAATCCATCAG
CTTCTGAGTAGTCATAGTCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCACAGGGGAAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAACCTCAAATCCATCAG
```

```
********** ** ********************************************************************************************************

CACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCAATCCAGGATCGATCCAGATCTTAAAGCTTACTAGTGGCTGAGCCTATAAAGCCTGCAGCCATG
CACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCAATCCAGGATCGATCCAGATCTTAA-GCTTACTAGTGGCTGAGCCTATAAAGCCTGCAGCCATG
************************************************************************** **********************************

GAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCC
GAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCACCATTGCC
************************************************************************************************************************

TTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGC
TTCACTGATGCTCACATTGAGGTGGACATCACCTATGCTGAATACTTTGAGATGTCTGTGAGGCTGGCAGAAGCCATGAAAAGATATGGACTGAACACCAACCACAGGATTGTGGTGTGC
************************************************************************************************************************

TCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGC
TCTGAGAACTCTCTCCAGTTCTTCATGCCTGTGTTAGGAGCCCTGTTCATTGGAGTGGCTGTGGCCCCTGCCAATGACATCTACAATGAGAGAGAGCTCCTGAACAGCATGGGCATCAGC
************************************************************************************************************************

CAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAAGACTGACTACCAGGGCTTC
CAGCCAACTGTGGTCTTTGTGAGCAAGAAGGGCCTGCAAAAGATCCTGAATGTGCAGAAGAAGCTGCCCATCATCCAGAAGATCATCATCATGGACAGCAAGACTGACTACCAGGGCTTC
************************************************************************************************************************

CAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGACTTTGTGCCTGAGAGCTTTGACAGGGACAAGACCATTGCTCTGATTATGAACAGCTCTGGCTCC
CAGAGCATGTATACCTTTGTGACCAGCCACTTACCCCCTGGCTTCAATGAGTATGACTTTGTGCCTGANAGCTTTGACAGGGACAAGACCATTGCTCTGATTATGAACAGCTCTGGCTCC
************************************************************** ********************************************

ACTGGACTGCCCAAAGGTG----------------------------------------
ACTGGNCTGCCCAAAGGTGTGGCTCTGCCCCACAGAACTGCTTGTGTGAGATTCAGCCAT
***** ************
```

pCpGl-exon2-R (mut)

```
                                                                new Hindiii
---------------------------------TGGATC--ACATCTCACAGGCCATATTTTTGCTTTATAAAAGGGAAGCTTAAAAGGCAATCATCGACAATATAGCCAGGTATTTGTG
ACAGCAGATTAAAAGGAATTCCTGCAGGACTAGTGGATCCCACATCTCACAGGCCATATTTTTGCTTTATAAAAGGGAAGCTTAAAAGGGAATCATCGACAATATAGCCAGGTATTTGTG
                               ******  ********************************************** ****************************

GTTCGTTATCATACTCTTCTGAGTAGTCATAGTCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCACAGGGGAAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAA
GTTCGTTATCATACTCTTCTGAGTAGTCAYAGWCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCACAGGGGAAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAA
***************************** ** ************************************************************************************

CCTCAAATCCATCAGCACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCAATCCAGATCCAGATCTTAAGCTTAGTCCATGGAGGATGCCAAGAATAT
CCTCAAATCCATCAGCACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCAATCCAGATCCAGATCTTAAGCTTAGTCCATGGAGGATGCCAAGAATAT
************************************************************************************************************************

TAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCC
TAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAACTGCACA----------------------------
************************************************************
```

pCpGl-AREG-exon2-AREG-promoter (mut)

```
GCAAACAGCAGATTAAAAGGAATTCCTGCAGGACTAGTGGATCGATCCTGGATTGGACCTCAATGACACCTACTCTGGGAAGCGTGAACCATTTTCTGGGGACCACAGTGCTGATGGATT
-------------------------------------TGCATCGATCCTGGATTGGACCTCAATGACACCTACTCTGGGAAGCGTGAACCATTTTCTGGGGACCACAGTGCTGATGGATT
                                  ** ********************************************************************************

TGAGGTTACCTCAAGAAGTGAGATGTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGWCTRTGACTACTCAGAAGAGTATGATAA
TGAGGTTACCTCAAGAAGTGAGATGTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGACTATGACTACTCAGAAGAGTATGATAA
************************************************************************************** ** ************************
                                    new Hindiii
CGAACCACAAATACCTGGCTATATTGTCGATGATTGCCTTTTAAGCTTCCCTTTTATAAAGCAAAAATATGGCCTGTGAGATGTGGAAGCTGATCCGGCTTGCAGCTAGAGGTTTAAGTT
CGAACCACAAATACCTGGCTATATTGTCGATGATTGCCTTTTAAGCTTCCCTTTTATAAAGCAAAAATATGGCCTGTGAGATGTGGAAGCTGATCCGGCTTGCAGCTAGAGGTTTAAGTT
************************************************************************************************************************

CCAYTTCCTCTCAGCGAATCCTTACGCACGAGGGAGGCGGGGCGTGTGTCCTCCGCGCGTGGTTTTCGGGTRGCACCTTCTGGGGCGCCGCCTGCCTCCACCCACGGCCGGGCCTTGACG
CCACTTCCTCTCAGCGAATCCTTACGCACGAGGGAGGCGGGGCGTGTGTCCTCCGCGCGTGGTTTTCGGGTAGCACCTTCTGGGGCGCCGCCTGCCTCCACCCACGGCCGGGCCTTGACG
*** ********************************************************** ********************************************

TCATGGGCTGCGGCCCCCTCCCGGCTGAGCCTATAAAGCGGCAGGTGCGCGCCGCCCTACAGACGTTCGCACACCTGGGTGCCAGCGCCCCAGAGGTCCCGGGACAGCCCGAGGCGCCGC
TCATGGGCTGCGGCCCCCTCCCGGCTGAGCCTATAAAGCGGCAGGTGCGCGCCGCCCTACAGACGTTCGCACACCTGGGTGCCAGCGCCCCAGAGGTCCCGGGACAGCCCGAGGCGCCGC
************************************************************************************************************************

GCCCGCCGCCCCGAGCTCCCCAAGCCTTCGAGAGCGGCGCACMCTCCCGGTCTCCACTCGCTCTTCCAACACCCGCTCGTTTTGGCGGCAGCTCGTGTCCCAGAGACCGAGTTGCCCCAG
GCCCGCCGCCCCGAGCTCCCCAAGCCTTCGAGAGCGGCGCACACTCCCGGTCTCCACTCGCTCTTCCAACACCCGCTCGTTTTGGCGGCAGCTCGTGTCCCAGAGACCGAGTTGCCCCAG
****************************************** ****************************************************************************

AGACCGAGACGCCGCCGCYGCGAAGGACCAATGAGAGCCCCGCTGCTACCGCCGGCGCCRGTGGTGCTGTCGCTCTTGATACTCGGCTCAGGTGAGGATTCACCGGCGCTGAACTGCTGG
AGACCGAGACGCCGCCGCTGCGAAGGACCAATGAGAGCCCCGCTGCTACCGCCGGCGCCGGTGGTGCTGTCGCTCTTGATACTCGGCTCAGGTGAGGATTCACCGGCGCTGAACTGCTGG
****************** ****************************************** ****************************************************

GCTCTCCTCCCATGGCAGGTTTTGCCTCTTGAGTAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAA
GCTCTCCTCCCATGGCAGGTTTTGCCTCTTGAGTAAGCTTAGTCCATGGAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCTACCCTCTGGAAGATGGCACTGCTGGTGAGCAA
************************************************************************************************************************
```

```
CTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCTGGCA
CTGCACAAGGCCATGAAGAGGTATGCCCTGGTCCCT----
************************************
```

pCpGl-AREG-exon2-R-AREG-promoter (mut)

```
                                                                    new HINDiii
-------------------------------------GCATCAGCTTC-ACATCTCACAGGCCATATTTTTGCTTTATAAAAGGGAAGCTTAAAAGGCAATCATCGACAATATAGCCAG
GCAAACAGCAGATTAAAAGGAATTCCTGCAGGACTAGTGGATCAGCTTCCACATCTCACAGGCCATATTTTTGCTTTATAAAAGGGAAGCTTAAAAGGGAATCATCGACAATATAGCCAG
                                  * ********* ********************************************** **********************

GTATTTGTGGTTCGTTATCATACTCTTCTGAGTAGTCATAGTCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCAC----------------------------------
GTATTTGTGGTTCGTTATCATACTCTTCTGAGTAGTCAYAGWCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCACAGGGGAAATCTCACTCCCTGAAGACATCTCACTTC
************************************ ** ****************************************
```

pCpGl-AREG-exon2-R-Tata (mut)

```
                                              new Hindiii
-------------TGGATCAGCTTC-ACATCTCACAGGCCATATTTTTGCTTTATAAAAGGGAAGCTTAAAAGGGAATCATCGACAATATAGCCAGGTATTTGTGGTTCGTTATCATACT
CCTGCAGGACTAGTGGATCAGCTTCCACATCTCACAGGCCATATTTTTGCTTTATAAAAGGGAAGCTTAAAAGGGAATCATCGACAATATAGCCAGGTATTTGTGGTTCGTTATCATACT
            ************ ********************************************** ****************************************

CTTCTGAGTAGTCATAGTCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCACAGGGGAAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAACCTCAAATCCATCAG
CTTCTGAGTAGTCAYAGWCGGCTCCCGAGGACGGTTCACTACTAGAAGGCATTTCACTCACAGGGGAAATCTCACTCCCTGAAGACATCTCACTTCTTGAGGTAACCTCAAATCCATCAG
************* ** ********************************************************************************************************
                                                                                                    TaTa-Box
CACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCAATCCAGGATCGATCCAGATCTTAA-GCTTACTAGTGGCTGAGCCTATAAAGCCTGCAGCCATG
CACTGTGGTCCCCAGAAAATGGTTCACGCTTCCCAGAGTAGGTGTCATTGAGGTCCAATCCAGGATCGATCCAGATCTTAAAGCTTACTAGTGGCTGAGCCTATAAAGCCTGCAGCCATG
************************************************************************** ************************************

GAGGATGCCAAGAATATTAAGAAAGGCCCTGCCCCATTCT
GAGGATGCCAAGAATATTA----------------------
******************
```

8.4 Example of lower results in promoter function analysis

Figure 67:

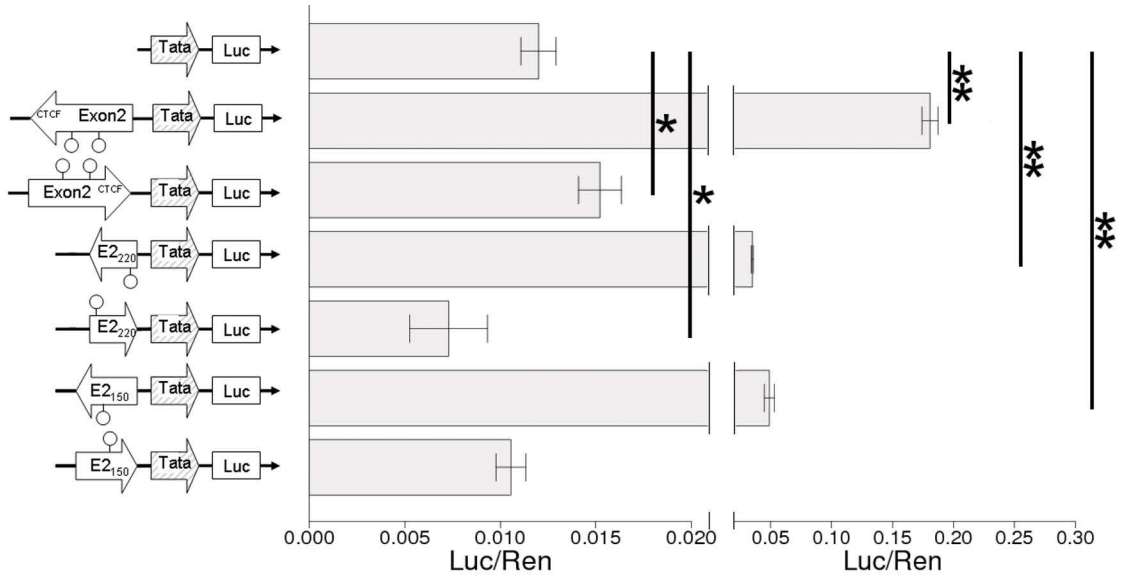


Figure 67: Verification of the t-test of the promoter function analysis: Figure 26 D was enlarged to show that some results, although they are not recognized in the original figure to be significant, have low non-overlapping error bars, which is an indication for significance in figures. Of course, the p-values calculated for these data proves significance.

8.5 *In vitro* methylation for promoter function analysis

Figure 68:

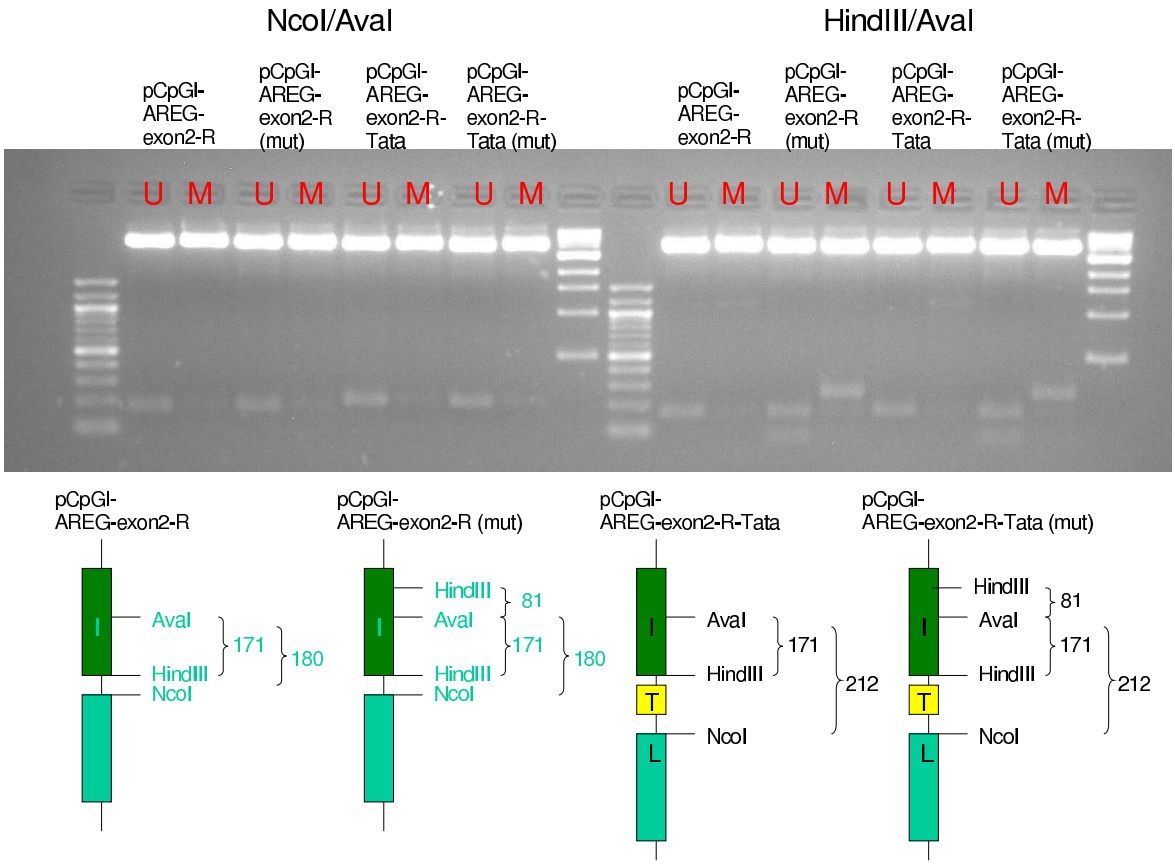


Figure 68: Analysis of *in vitro* methylation by restriction digest: Four plasmids were methylated as described in section 2.2.8 and afterwards either digested with NcoI/AvaI (left) or with HindIII/AvaI (right). Methylated CpGs and the mutation of the CTCF-binding site lead to different restriction bands which are shown in the schemes below the gel figure. AvaI is the methylation dependent restriction enzyme. In the agarose gel the unmethylated plasmid (U) is shown next to the methylated plasmid (M) to compare the restriction digest bands.

8.6 Expression plasmids for lentiviral transfection

Figure 69:

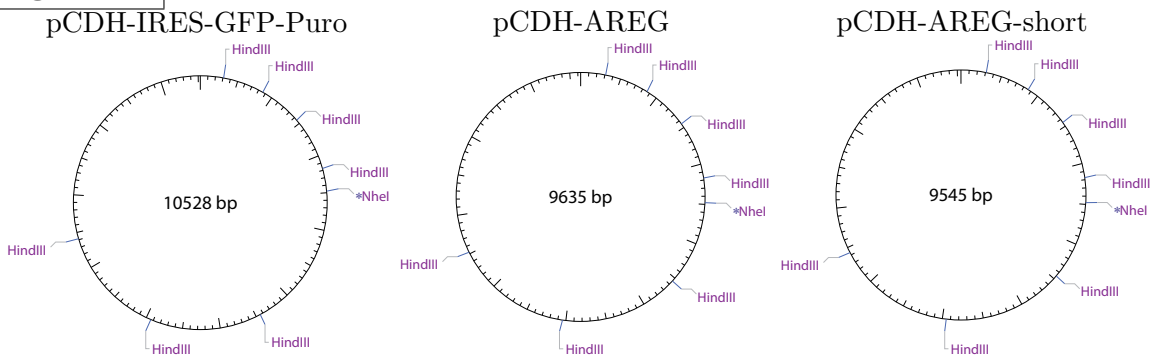


Figure 69: Restriction maps of expression plasmids: PCDH-IRES-GFP-Puro, pCDH-AREG and pCDH-AREG-short were digested *in silico* with HindIII and NheI using the webtool NEBcutter.[233]

Table 25: Linear fragments after restriction digest of expression plasmids with HindIII and NheI

No	Ends	pCDH-IRES-GFP-Puro	pCDH-AREG	pCDH-AREG-short
1	HindIII-HindIII	3424	3424	3424
2	**NheI-HindIII**	**1960**	**1067**	**977**
3	HindIII-HindIII	1496	1496	1496
4	HindIII-HindIII	1479	1479	1479
5	HindIII-HindIII	728	728	728
6	HindIII-HindIII	583	583	583
7	HindIII-HindIII	553	553	553
8	HindIII-NheI	305	305	305

Figure 70:

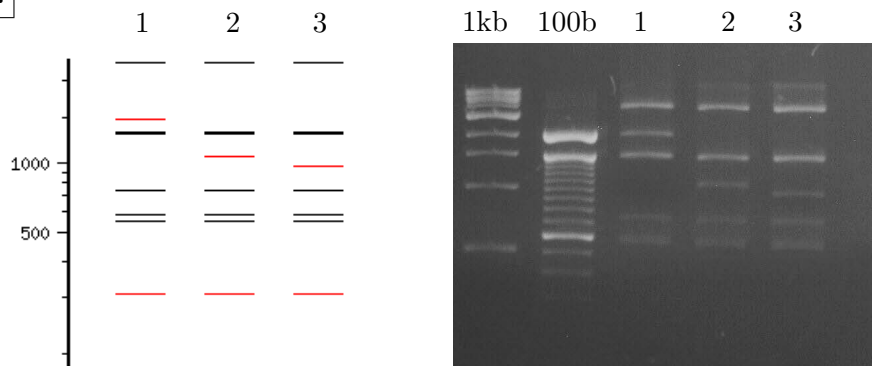


Figure 70: *In silico* VS *in vitro* digest of pCDH-plasmids: 1: PCDH-IRES-GFP-Puro, 2: pCDH-AREG and 3: pCDH-AREG-short were digested *in silico* using the webtool NEBcutter[233] and *in vitro* with HindIII and NheI.

8.7 Control digest for Northern blot probe generation

Figure 71:

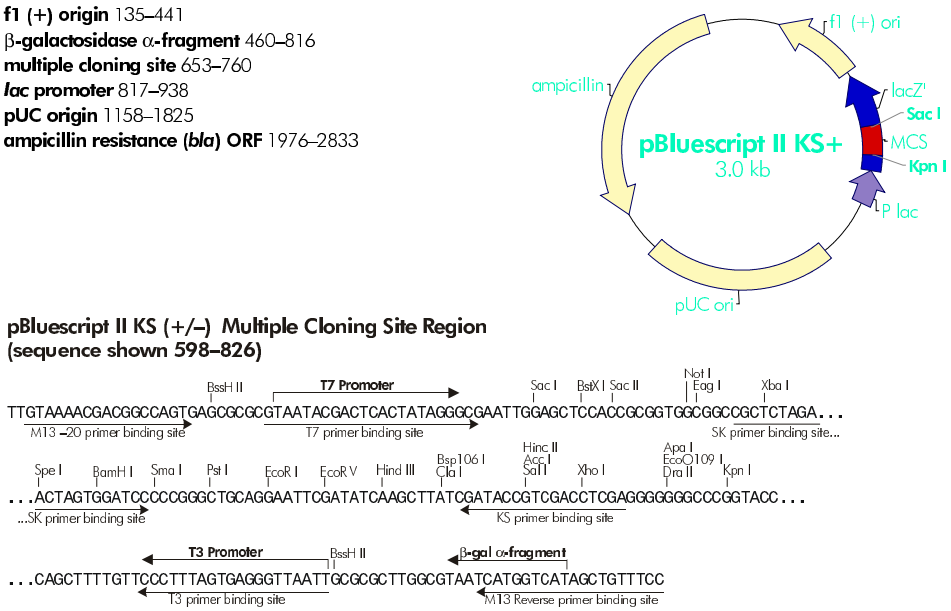


Figure 71: Plasmid map of pBluescript II KS +: Inserts were cloned into the plasmid pBluescript II KS + after digest with EcoRV for Northern blot probe generation. Plasmids were sequenced with the sequencing primer T7. Plasmids were linearized either with HindIII or EcoRI which depended on the orientation of the inserts.

Figure 72:

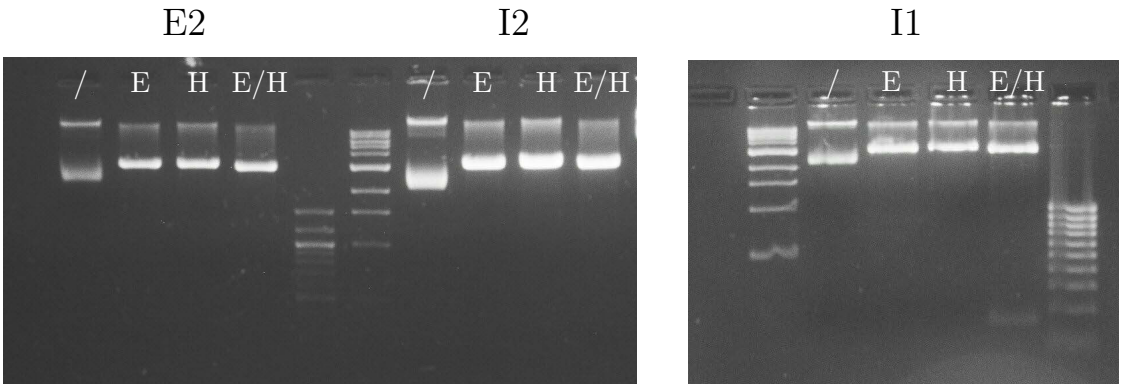


Figure 72: Control digest of pBluescript II KS + -derived plasmids: Plasmids were digested either with HindIII or EcoRI or with both to analyze if only one insert is located within the plasmid after cloning

8.8 Sequencing results of plasmids for Northern blot probe generation

pBluescript II KS + - E2
Insert sequence orientation: forward, Sequenzing primer: T7-forward

```
----------------------------------------NNNNGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGATCCACAGTGCTGATGGATTTGAGGTTACCT
GTAATACGACTCACTATAGGGCGAATTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTC---CCACAGTGCTGATGGATTTGAGGTTACCT
                                        ******************************************   ****************************

CAAGAAGTGAGATGTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGACTATGACTACTCAGAAGAGTATGATAACGAACCACAAA
CAAGAAGTGAGATGTCTTCAGGGAGTGAGATTTCCCCTGTGAGTGAAATGCCTTCTAGTAGTGAACCGTCCTCGGGAGCCGWCTRTGACTACTCAGAAGAGTATGATAACGAACCACAAA
************************************************************************** ** **********************************

TACCTGGCT---ATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACC-AGCTTTTGTTCCCTTTAGTGAGGGTTAATTTCGAGCTTGGCGTAATCATGGTCATAGCTGTT
TACCTGGCTGATATCAAGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTT
*********   ****************************************** **************************** ** ****************************
```

pBluescript II KS + - I1
Insert sequence orientation: forward, Sequenzing primer: T7-forward

```
----------------------------------------GGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGATTAGAAATGACTCAATCACAAACAATAGCA
GTAATACGACTCACTATAGGGCGAATTGGAGCTCCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGA-----AATGACTCAATCACAAACAATARCA
                                    ************************************************     ********************** **

CATGTGCAATAACTGCTTTGATGTCAAAAGATAAACTTTTCTACCTAACACAGCTGACTCGAAAGGCACCCTACTTTACCTTTTCTTTTCTTCCCTTATTCCCTCCCCTGCAG---ATCA
CATGTGCAATAACTGCTTTGATGTCAAAAGATAAACTTTTCTACCTAACACAGCTGACTCGAAAGGCACCCTACTTTACCTTTTCTTTTCTTCCTTTATTCCCTCCCCTGCAGGATATCA
************************************************************************************************** *****************   ****

AGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACC-AGCTTTTGTTCCCTTTAGTGAGGGTTAATTTCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGT
AGCTTATCGATACCGTCGACCTCGAGGGGGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGCTTGGCGTAATCATGGTCATAGCTGTTTCC---
************************************* **************************** ** ******************************
```

All inserts were sequenced correctly. The three to five bases ins/dels were identified to be the former EcoRV-site. Therefore, they flank the inserts. After the insert orientation was identified, EcoRI or HindIII digests were performed to linearize the plasmid for probe generation.

Table 26: Linearization of pBluescript II KS + -derived plasmids

PCR-prod.	Insert-orientation.	Restriction digest	RNA-Pol.	Probe
AREG exon 2	forward	HindIII	T7	E2-antisense
AREG intron 1	forward	HindIII	T7	I1-antisense
AREG exon 2	forward	EcoRI	T3	E2-sense

8.9 The effect of Zebularine on *AREG* expression

Figure 73:

120 h: **144 h:**

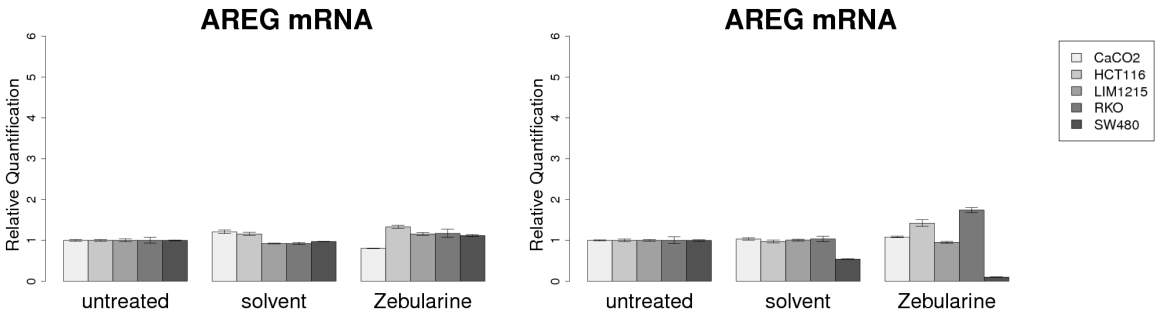


Figure 73: *AREG* mRNA expression after Zebularine treatment: Five cell lines were testet for their *AREG* mRNA expression after treatment with 100 µM Zebularine for 120 and 144 h. In each plot the relative quantification of the expression is shown compared to the untreated control.

Figure 74:

120 h: **144 h:**

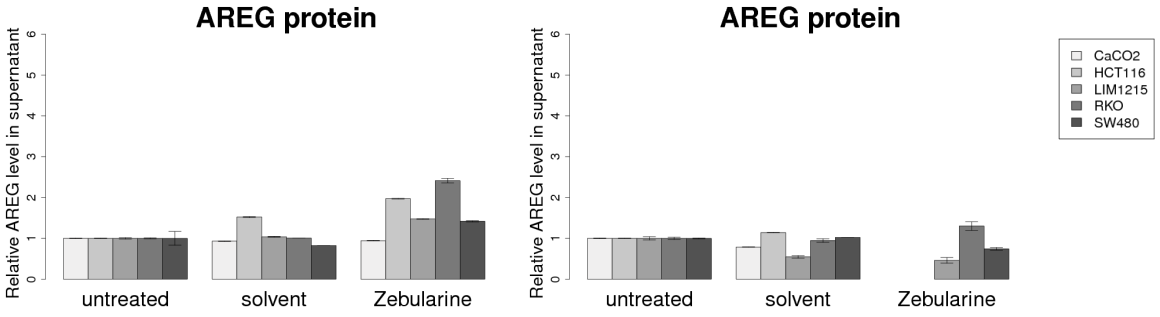


Figure 74: AREG protein expression after Zebularine treatment: Five cell lines were testet for their AREG protein amount in the supernatant after treatment with 100 µM Zebularine for 120 and 144 h. In both plots the relative AREG level is shown compared to the untreated control.

8.10 Betacellulin: gene expression VS methylation

Figure 75:

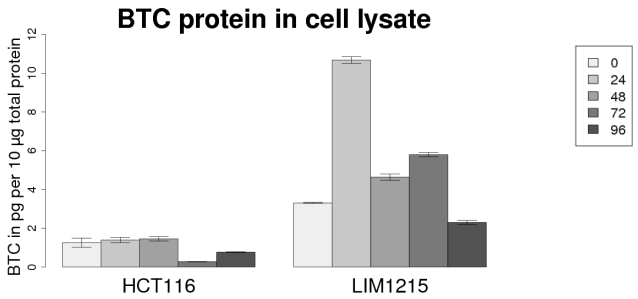


Figure 75: Protein expression of BTC in HCT116 and LIM1215: HCT116 and LIM1215 were tested for their BTC-protein expression in cell lysate by ELISA. Cells were seeded into 6-well plates as described in section 2.3.4 and cell lysates were isolated as described in section 2.5.1, starting 1 day after seeding (timepoint 0). BTC-ELISA was performed as described in section 2.5.2.

Figure 76:

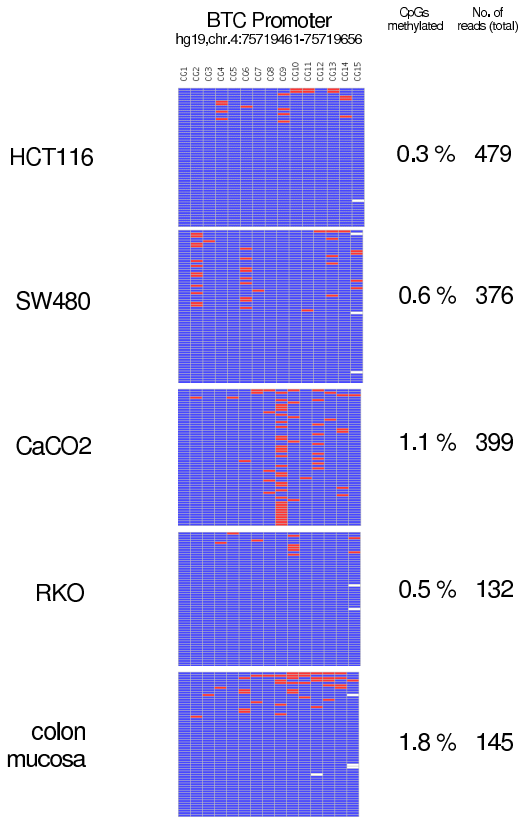


Figure 76: Methylation of the *BTC* promoters after 454 GS-FLX deep bisulfite amplicon sequencing: Four different cell lines and colon mucosa DNA as a negative control were tested for their *BTC* promoter methylation. Per sequenced sample, 40 to 50 reads, represented by the lines, are shown. CpG positions are represented by the columns. Genomic positions are defined by the first proximal and last distal CpG of the analysed amplicon. Averaged DNA methylation and total number of reads are given at the right of the respective map. Blue: the CpG is unmethylated, Red: the CpG is methylated, White: the methylation status is not analyzed.

Figure 77:

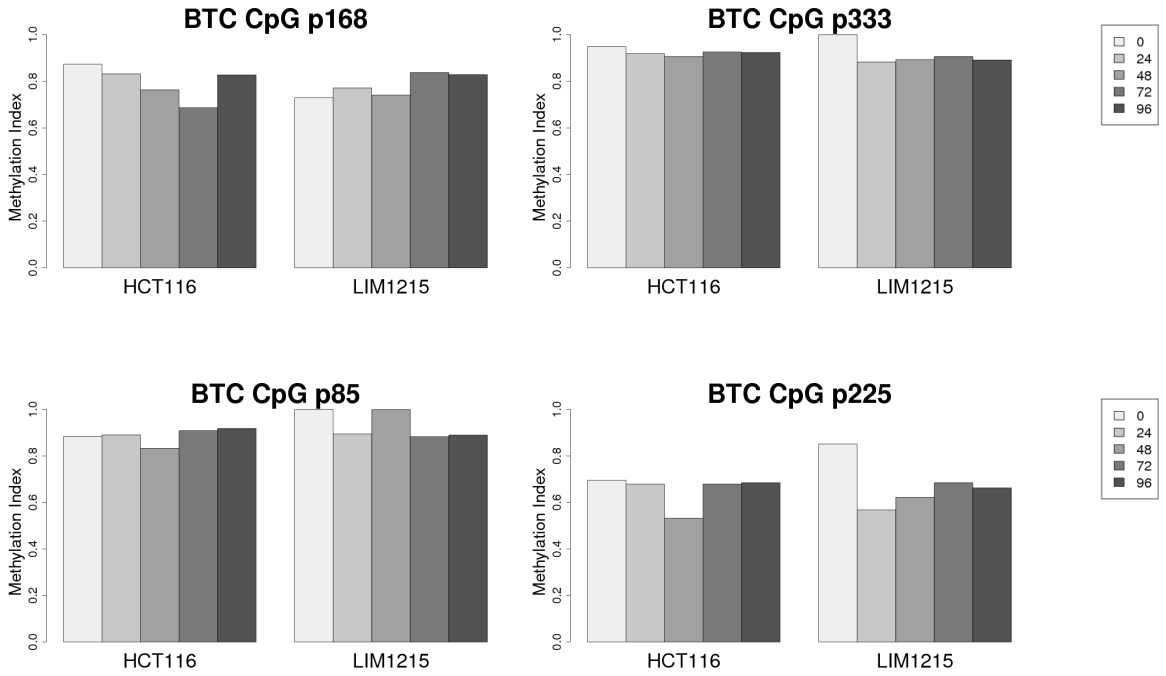


Figure 77: Methylation of four CpGs within the *BTC* gene: HCT116 and LIM1215 were tested for their methylation of intragenic CpGs within the BTC-gene. CpG p168 and CpG p333 are located within intron 2, CpG p85 and CpG p225 are located within intron 4. Methylation was determined by MsSNuPE experiments (see section 2.12.1. A methylation index of 1 means fully methylated, whereas a methylation index of 0 means fully unmethylated.

8.11 Treatment of patient-derived xenografts with Cetuximab

Table 27: Xenograft experiment: Treatment of *KRAS*-wildtype xenografts with Cetuximab

ID	type	TNM-Status	KRAS (12+13)	CT AREG	pgAREG/mg total protein	Cetuximab
7888	metastasisof rectal carcinoma (liver)	yrepTxNxM1R0LxVxGx	**wt**	3.32	8280	+
5735	colon carcinoma	T3N2G2 M0	**wt**	4.20	7180	++
7475	metastasis of rectal carcinoma (lung)	M1	**wt**	1.62	6220	-
7596	metastasis of rectal carcinoma (lung)	M1	**wt**	3.46	4420	+
5676	colon carcinoma (rectum)	T3aN0G2L0V1R0M0	**wt**	5.72	3060	++++
5771	colon carcinoma	T3N1G2LxV0R0M0	**wt**	3.62	1368	+
7818	metastasis of colon carcinoma (lung)	rpTxNxM1R0LxVxG2	**wt**	6.97	1060	-
5854	colon carcinoma	T4N2G3LXV1R0M0 (Black ed.)	**wt**	7.48	613	-
7553 / B	metastasis of colon carcinoma (lung)	M1	**wt**	11.58	102	-

Table 28: Xenograft experiment: Treatment of *KRAS*-mutant xenografts with Cetuximab

ID	type	TNM-Status	KRAS (12+13)	CT AREG	pgAREG/mg total protein	Cetuximab
5682	colon carcinoma (rectum)	T4bN2G2L0V1R0M0	**mut**	NA	6430	-
5679	colon carcinoma	T3N2G3L1V1R0M1 (liver)	**mut**	5.26	1638	-
7835	metastasis of colon carcinoma (lung)	rpTxNxM1R0LxVxG2	**mut**	4.99	1348	++
7809	metastasis of colon carcinoma (liver)	yrp TxNxM1R0LxVxG3	**mut**	5.32	1112	+
5896	colon carcinoma	T2N0G2L0V0R0 M0	**mut**	4.24	1027	-
5736	colon carcinoma (rectum)	T4N0G2L0V0R0M0	**mut**	NA	774	-
5841	colon carcinoma (rectum)	T3N2G3L1V1R0M1 (liver)	**mut**	8.15	507	-
7689	metastasis of rectal carcinoma (lung)	M1	**mut**	7.14	210	-

8.12 *AREG* intragenic methylation in 4 microdisected human cancer tissues

Table 29: *AREG* intragenic methylation in microdisected tumor samples

No.	ID	tissue	region	AREG CpG p150	AREG CpG p220
1	4548/11	Normal	Epithel	0.178	0.255
2			Stroma	0	0
3		Tumor	Epithel	0	0.237
4			Stroma	0.328	0.373
5	7135/11	Normal	Epithel	0.851	1
6			Stroma	0.779	1
7		Tumor	Epithel	0	0
8			Stroma	1	1
9	42460/11	Normal	Epithel	NA	NA
10			Stroma	0.805	0.561
11		Tumor	Epithel	0	1
12			Stroma	1	1
13	43604/11	Normal	Epithel	0.803	1
14			Stroma	0.723	0.793
15		Tumor	Epithel	0	0
16			Stroma	1	1
17	43467/11	Normal	Epithel	1	1
18			Stroma	0.115	0.364
19		Tumor	Epithel	0.242	0
20			Stroma	0.593	0.547

8.13 *AREG* and *EREG* intragenic methylation in 24 microdisected human cancer tissues

Figure 78:

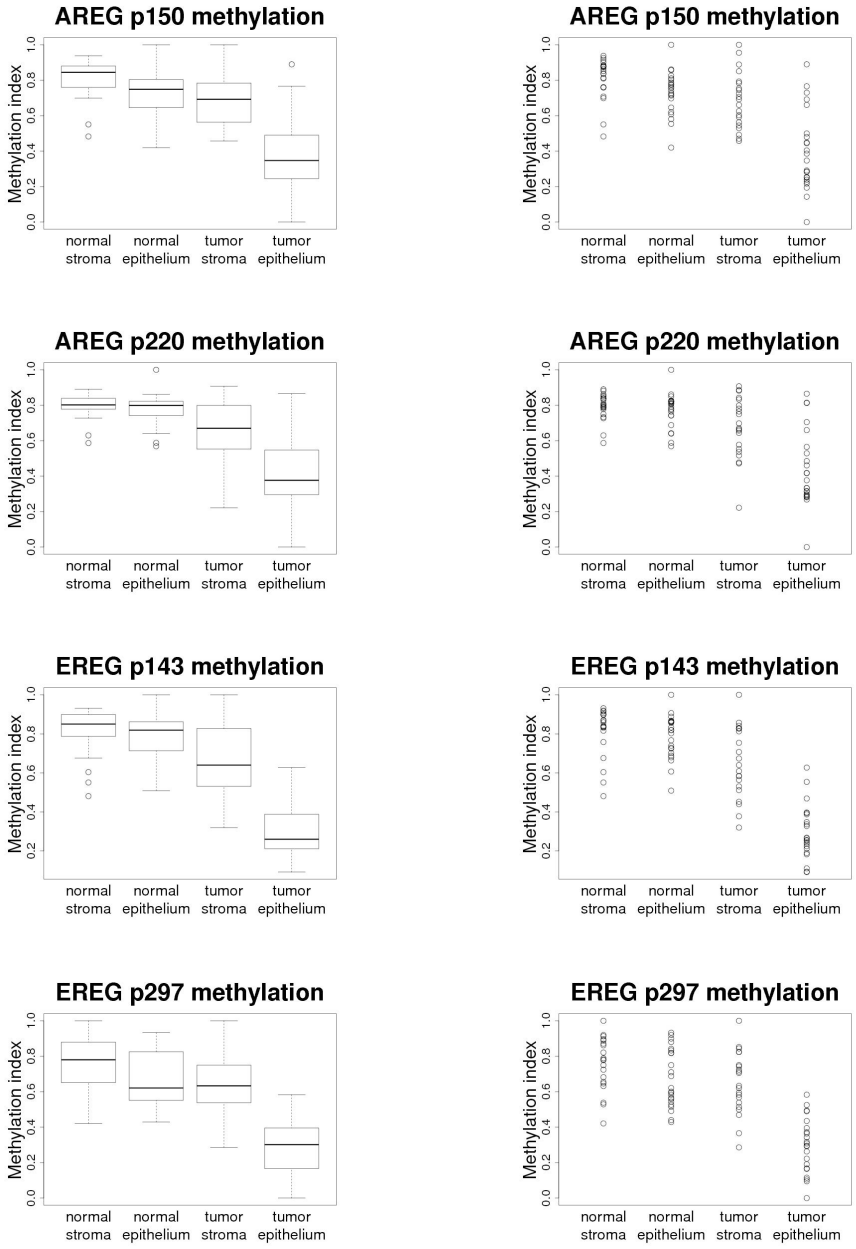


Figure 78: *AREG* and *EREG* intragenic CpG methylation in microdisected human cancer tissues: 24 human colorectal cancer tissues were microdisected into four parts (normal stroma, normal epithel, tumor stroma and tumor epithel) AREG and EREG intragenic methylation was measured by msSNuPE (see section 2.12.1) The Methylation Indices of each sample were either presented as boxplot (left) or directly (right).

8.14 *AREG* mRNA expression after Erlotinib and Gefitinib treatment

Figure 79:

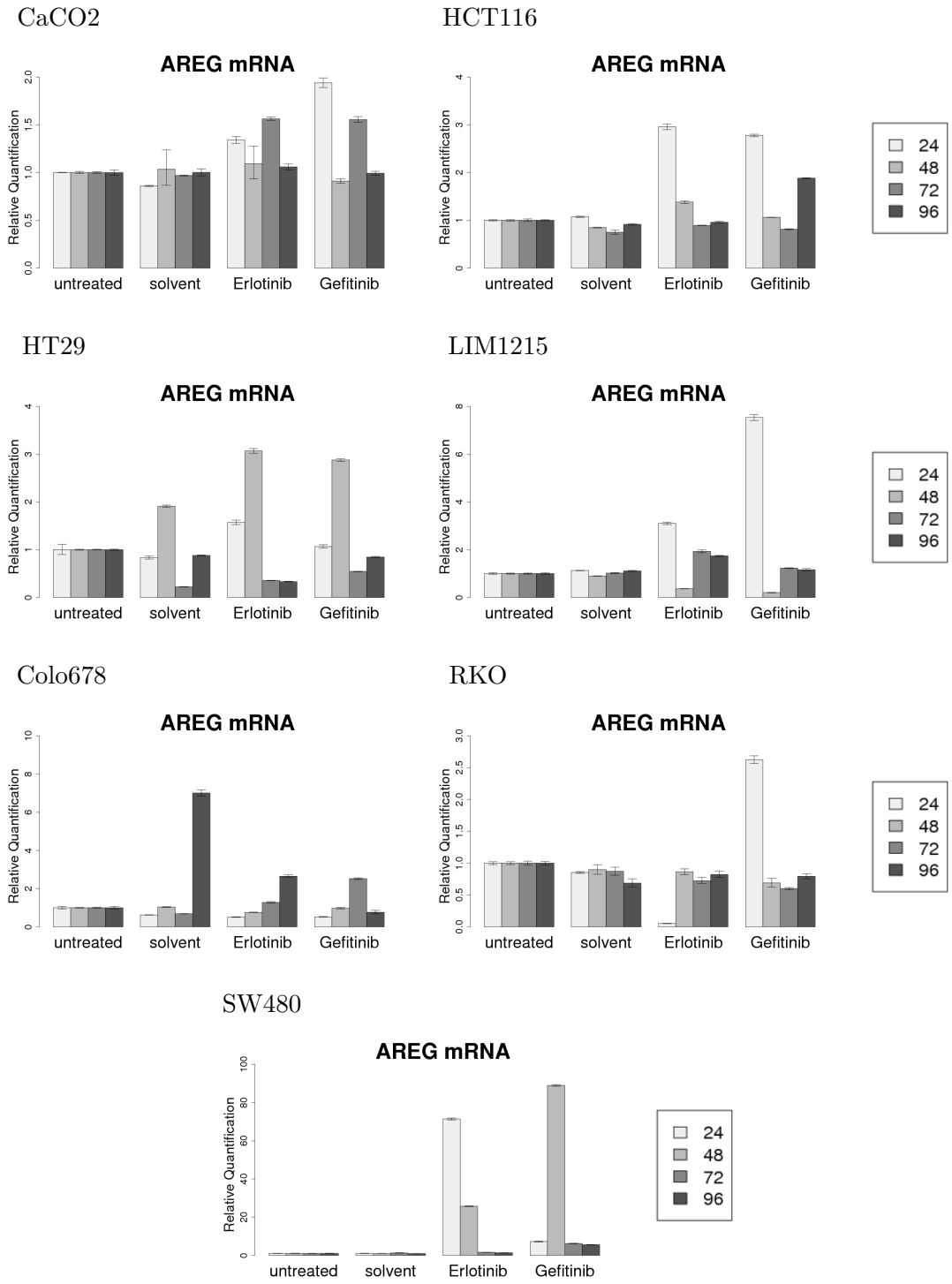


Figure 79: *AREG* mRNA expression after EGFR inhibitor treatment in different colorectal cancer cells: Seven colorectal cancer cell lines were tested for their *AREG* mRNA expression after treatment with 10 µM Erlotinib or 10 µM Gefitinib. The relative quantification of the mRNA expression is shown for four different timepoints. solvent: DMSO

9 Publications

Articles

Bormann, F., Sers, C., Seliger, B., Handke, D., Bergmann, T., Seibt, S., Lehrach, H. and Dahl, A. (2011). Methylation-specific ligation detection reaction (msLDR): a new approach for multiplex evaluation of methylation patterns. *Mol Genet Genomics* 286:279-291.

Bormann, F., Dahl, A. and Sers, C. (2012). MsLDR-creator: a web service to design msLDR assays. *Mol Genet Genomics* 287:273-274.

Patents

provisorische Patentanmeldung: Europische Patentanmeldung, "Nachweis von DNA Methylierung in den Genen AREG, EREG und BTC" CH594/2011

Poster presentations

Dahl, A., Bormann, F., Schwartz, R., Lange, M., Lehrach, H. and Nyarsik, L. (2007). SNP fine mapping on a miniaturized platform using TaqMan assays on a nanoliter scale, Day of Science (MPI for Molecular Genetics), Berlin.

Bormann, F., Lange, M., Nyarsik, L., Lehrach, A. and Dahl, A. (2008). Development of a contact DNA transfer device for the µPCR-chip platform, Day of Science (MPI for Molecular Genetics), Berlin.

Bormann, F., Kelm, A.S., Dietel, M., Tierling, S., Walter, J., Rivera-Markelova, M., Fichtner, I., Schäfer, R., Weichert, W. and Sers, C. (2010). Regulation of the EGFR ligands AREG, EREG and BTC through DNA methylation-dependent mechanisms in colorectal cancer, Transcription, chromatin structure and DNA repair in development and differentiation, Essen.

Bormann, F., Kelm, A.S., Dietel, M., Tierling, S., Walter, J., Rivera-Markelova, M., Fichtner, I., Schäfer, R., Weichert, W. and Sers, C. (2011). Epigenetic regulation of the EGFR ligand Amphiregulin and its role in predicting the outcome of EGFR-targeted therapies in colorectal cancer, 16th International AEK Cancer Congress, Düsseldorf.

Bormann, F., Tierling, S., Walter, J. and Sers, C. (2012). Epigenetic regulation of Amphiregulin and Epiregulin and their impact on the outcome of EGFR targeted therapeutics in colorectal cancer cell lines, 22. EACR congress, Barcelona, Spain.

Bormann, F., Bartels, S., Tierling, S., Walter, J., Sers, C. (2013). Epigenetic regulation of Amphiregulin and Epiregulin and their impact on the outcome of EGFR targeted therapeutics in colorectal cancer cell lines, Keystone Symposium: Epigenetic Marks and Cancer Drugs, Santa Fe (NM), USA.

Berlin, den

_____

(Felix Bormann)

SELBSTSTÄNDIGKEITSERKLÄRUNG

Hiermit erkläre ich, dass ich die vorliegende Arbeit eigenständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Stellen, die wörtlich oder sinngemäß aus Quellen entnommen wurden, sind mit der Quellenangabe versehen. Ich versichere, dass ich mich nicht anderweitig um einen Doktorgrad beworben habe oder einen entsprechenden Doktortitel besitze. Die Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät I der Humboldt-Universität zu Berlin habe ich gelesen und akzeptiert.

Berlin, den

_____

(Felix Bormann)